PRP

Vol. 48 OCTOBER, 1957 No. 6

The Journal of Educational Psychology

Desk

UNIVERSITY OF MICHIGAN
DEC 2 1957
PERIODICAL READING ROOM

CONTENTS

The Learning of Motor Skills as Influenced by Knowledge of Mechanical Principles 321
FRANCES M. COLVILLE

Student vs. Instructor Design of Study Method 328
SLATER E. NEWMAN

Retroactive Inhibition and Facilitation in the Learning of School Materials 334
DAVID P. AUSUBEL, LILLIAN CUKIER ROBBINS, AND ELIAS BLAKE, JR.

The Identification of Number Anxiety in a College Population 344
RALPH MASON DREGER AND LEWIS R. AIKEN, JR.

The Relation of Departure Time and Retension to Academic Achievement 352
ABRAM M. BARCH

Differentiating Characteristics of a Group of Students Having Psychological Problems 359
GEORGE DOUGLAS MAYO

Comparative Effectiveness of Televised and Large Classes and of Small Sections 371
LAURENCE SIEGEL AND F. G. MACOMBER

Book Reviews 383

Published monthly except June to September
$7.00 a year in U. S. and Pan America; Canada, $7.20; other countries, $7.40.
Single issues $1.20

AMERICAN PSYCHOLOGICAL ASSOCIATION

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879.

THE JOURNAL OF Educational Psychology

Established 1910

BOARD OF EDITORS*

WM. CLARK TROW, *Chairman*
University of Michigan

STEPHEN M. COREY
Teachers College
Columbia University

J. B. STROUD
State University of Iowa

*Appointed by the American Psychological Association.

EDITORIAL CONSULTANTS

PAUL BLOMMERS
CHESTER W. HARRIS
HAROLD E. JONES
H. H. REMMERS
PERCIVAL M. SYMONDS
MILES A. TINKER
ALEXANDER G. WESMAN
PAUL A. WITTY

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. THE JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

AMERICAN PSYCHOLOGICAL ASSOCIATION
1333 Sixteenth Street, N. W.
Washington 6, D. C.

Copyright 1957 by the American Psychological Association, Inc.

UNIVERSITY OF MICHIGAN LIBRARIES

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume 48 October, 1957 Number 6

THE LEARNING OF MOTOR SKILLS AS INFLUENCED BY KNOWLEDGE OF MECHANICAL PRINCIPLES

FRANCES M. COLVILLE

Santa Barbara College, University of California

Many general principles of mechanics are relevant to the teaching, learning, and performance of activities included in the physical education curriculum. These principles may describe the motion of objects such as balls and racquets, movement of the body itself, or a combination of both.

The question has been raised as to whether knowledge of these general principles and an understanding of their application to these activities will facilitate the learning and improve the performance of pupils in physical education classes. Also of interest are questions concerning the influence of knowledge of these principles as applied to one activity upon subsequent learning of other activities to which the same principles apply. Results of a few experiments suggest that a pupil who understands a principle related to one skill may master a related skill more readily than the pupil whose experience has been restricted to specific instruction in technique without explanation of pertinent principles.

With reference to the teaching of a motor skill in which a specific principle of mechanics is involved, then, two questions have been raised:

1. What is the effect of knowledge of a principle upon immediate learning of a skill to which the principle applies?
2. What is the effect of knowledge of a principle learned in relation to one skill upon subsequent learning of a different or more complicated skill to which the same principle is applicable?

Reported investigations bearing on the acquisition and transfer of principles are limited, and although most authorities agree that

general principles are probably transferable, little strong experimental evidence is available to support such a claim. Apparently, little or no experimentation has been done in the field of physical education on the effect of knowledge of principles of mechanics in learning situations involving body movement and large muscle activities.

Judd (**7**), and Hendrickson and Schroeder (**4**) reported that knowledge of the principle of refraction was beneficial in dart throwing and rifle shooting at a target under water. Cox (**2**) found that knowledge of principles involved in manipulating electrical equipment facilitated the assembling and stripping of such equipment. Other experimenters have reported similar results from investigations involving mathematics, mental games, spelling, card tricks, and mechanical puzzles (**13, 11, 15, 10, 14, 8, 2**). In contrast to these findings are a number of investigations in which the method of including instruction in the principles involved did not appear to be more effective. In studying the effect of drill versus the learning of generalizations related to addition and subtraction, Olander (**9**) found that the two methods appeared to be equally effective. Babitz and Keys (**1**) found similar results with college chemistry classes, and Hendrix (**6**) reported only one difference, significant at the 12% level, in her experiment with teaching algebraic generalizations.

With the exception of the experiments by Judd, Hendrickson and Schroeder, and Cox, all of the reported investigations have dealt with various types of learning other than motor learning, and only Judd's experiment involved skills requiring the coordinated use of large muscles of the body as well as fine coordinations of the fingers. Judd, and Hendrickson and Schroeder, while reporting a difference in favor of instruction in principles, included this instruction in such a way that additional time was allotted to the group receiving the instruction. Thus, the experimental groups had the advantage of additional time as well as additional information. Furthermore, Judd supplied no statistical analysis of his early experiment, and the other differences reported were significant at a statistical level below that which is usually considered acceptable.

In the present investigation, an attempt was made to control these ambiguities by using an experimental design which insured

UNIVERSITY OF MICHIGAN LIBRARIES

that both the experimental and the control groups spent equal amounts of time in learning and performing and by considering only those differences which were significant at the 5% level of confidence or above.

Method

Purpose of the Study

It was the purpose of this study to investigate certain questions related to the teaching of physical education activities in which specific principles of mechanics are involved. This problem was approached by: (*a*) selecting three principles of mechanics which are pertinent to motor skills; (*b*) selecting three motor skills, each of which utilizes one of the principles; (*c*) establishing for each skill two comparable groups of *S*s, one of which was taught without reference to the principle involved, and the other of which was taught to understand and apply the principle; (*d*) comparing for each skill the performance of the two groups, one of which spent the entire time practicing the skill without reference to the principle involved, and the other of which spent part of the time practicing the skill and part of the time in learning the principle; (*e*) comparing for each skill the performance of the two groups in a similar or more complicated form of the skill to which the same general principle applies.

Selected Principles and Skills

The selected principles and the skills in which they are utilized were as follows:

Principle I: The angle of incidence is approximately equal to the angle of reflection. Skill: Rolling a ball against a surface, or surfaces, from which it would rebound.

Principle II: In stopping a moving object, the force opposing the momentum must be equal to the force of the momentum, and if the object is to be caught, this momentum must be dissipated by reducing the resistance of the catching surface. Skill: Catching a tennis ball in a lacrosse stick and catching a badminton bird on a tennis racquet.

Principle III: An object set in forward motion through the air by an external force is acted upon by the force of the momentum and by gravital acceleration. Skill: Archery.

Procedure

Three parallel experiments were devised, each of which was designed to investigate performance of one of the three skills. The *S*s for all three experiments were undergraduate women students at the University of Southern California. There were 36 *S*s in the Ball Rolling experiment, 40 in the Catching experiment, and 42 in the Archery experiment. (The *S*s in the Archery experiment included some men whose scores were treated separately from those of the women.) For each of the experiments the *S*s were divided into two groups, one of which spent the entire time learning and practicing the skill, and the other of which spent part of the time learning about the principle and the rest of the time learning and practicing the skill. Each experiment consisted of two tests. During the First test the *S*s of the nonprinciple group learned the skill and practiced. The *S*s of the principle group learned the principle, and the skill, and practiced. The total amount of time spent by both groups was the same. During the Second test both groups practiced a similar or more complicated skill. Each group spent the same amount of time and no mention was made of the principle involved.

The Ball Rolling experiment. The Ball Rolling experiment was designed to illustrate the principle describing the rebound of a ball. The appratus consisted of a modified pinball machine plunger, a squash ball, a felt-covered table surface four feet wide and six feet long, and a target made of colored construction paper. The overall size of the target and each of its scoring divisions was determined by recording the hits made by a group of comparable *S*s who were not used in the experiment. These *S*s rolled the ball so that it would rebound from one side of the table as close as possible to the center of the adjacent side which was marked with a square of white paper. The rest of the side was covered with brown wrapping paper, and spots on this paper to which the balls rolled were marked in red pencil. This scoring sheet provided the basis for establishing a six-standard deviation scale, and from this scale the scoring areas of the target were derived.

The *S*s were asked during the First test to roll the ball against one side of the table in such a manner that it would rebound to the target placed against the adjacent side. For the Second test they were asked to roll the ball so that it would rebound from two adjacent sides to the target set at right angles on the third side.

UNIVERSITY OF MICHIGAN LIBRARIES

The Catching experiment. The Catching experiment consisted of two tests each involving the skill of catching a moving object. The First test involved catching a tennis ball in a lacrosse stick. For the Second test, they caught a badminton bird on a tennis racquet. The apparatus for this experiment, in addition to balls, birds, crosse, and racquet, included a homemade device for projecting the tennis ball and the badminton bird. This device proved to have high reliability.

The Archery experiment. The Archery experiment was planned to illustrate the principle of gravital acceleration and its action on arrow flight and the relation of line of flight to line of sight. The First test involved shooting from 20 yards. The Second test included ends shot from 30 yards, from 40 yards, and a Junior Columbia Round.

Analysis of data. In each experiment, the reliability of the scores was established by use of the split-half technique, the odd scores being correlated with the even scores. Use of the Spearman-Brown Prophecy Formula for predicting reliability of the whole test resulted in coefficients which ranged from .81 to .94 and thus it was felt that these scores were sufficiently reliable to permit further analysis and comparison.

All of the data were analyzed by means of analysis of variance using Edwards' (**3**) technique for repeated measurements of the same subjects. This technique tests differences between methods, between trials, and between the interaction of trials and methods. In each experiment the initial ability of the two groups, as evidenced by preliminary trials, was tested and found not to differ significantly.

Findings

In the Ball Rolling experiment, no significant difference in the performance levels of the two groups was found on either test. A difference between trials significant at the 1% level of confidence was found on both tests.

In the Catching experiment, no significant differences were found in performance levels, and a difference between trials significant at the 5% level of confidence was found only on the First test.

In the Archery experiment no significant difference in performance levels was found. A difference significant at the 1% level for men and at the 5% level for girls was found on the Second test at all distances except forty yards.

The following general findings were noted:

1. A significant amount of learning took place under both methods of instruction.

2. This learning was not only similar in amount but also in pattern. Exceptions occurred in the Second test of the Catching experiment where no significant increase in scores occurred in either group, and in the Junior Columbia Round for men in the Archery experiment, where both groups increased in scoring ability, but the group which had been taught the principle was significantly better at twenty yards.

These findings are in general agreement with those of some published investigations and are opposed to those of some others. However, most of the reported experiments have dealt with various types of learning other than motor learning. The ability to use understood principles in performing a motor skill presents a complex problem since more than one type of learning is involved. As Ragsdale (**12**) has pointed out, a learner may understand a principle when applied to inanimate objects, but he may not understand and be able to apply it to his own movements. In general, it seems that the findings of this study support Ragsdale's observation, at least in the initial stages of acquiring an unfamiliar motor skill.

Conclusions

Within the limitations of the three parallel experiments which constitute the present investigation, there is no evidence:

1. That instruction concerning mechanical principles utilized in the performance of a motor skill facilitates the initial learning of the skill to any greater extent than an equivalent amount of time spent in practicing the skill.

2. That such knowledge facilitates subsequent learning as evidenced in the performance of a similar or more complicated skill to which the same principle is applicable.

However, since it appears that some part of the learning period may be devoted to instruction concerning general principles without detriment to the motor learning of the students, it would seem desirable to include such instruction in order to provide this additional opportunity for acquiring some related knowledge about principles of mechanics and the application of forces.

UNIVERSITY OF MICHIGAN LIBRARIES

References

1. Babitz, M., & Keys, N. Experiment in teaching pupils to apply scientific principles. *Sci. Educ.*, 1939, **23**, 367–370.
2. Cox, J. W. Some experiments on formal training in the acquisition of skill. *Brit. J. Psychol.*, 1933, **24**, 67–87.
3. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
4. Hendrickson, G., & Schroeder, W. H. Transfer of training in learning to hit a submerged target. *J. Educ. Psychol.*, 1941, **32**, 205–213.
5. Hendrickson, G. & Schroeder, W. H. Transfer of training. In R. G. Kuhlen & G. G. Thompson (Eds.), *Psychological studies of human development.* New York: Appleton-Century-Crofts, 1952.
6. Hendrix, Gertrude. A new clue to transfer of training. *Elem. Sch. J.*, 1947, **48**, 197–208.
7. Judd, C. H. Special training and general intelligence. *Educ. Rev.*, 1908, **36**, 38–42.
8. Katona, G. *Organizing and memorizing.* New York: Columbia Univer. Press, 1940.
9. Olander, H. T. Transfer of learning in simple addition and subtraction. *Elem. Sch. J.*, 1931, **32**, 358–369; 427–437.
10. Overman, J. R. Experimental study of the effect of the method of instruction on transfer of training in arithmetic. *Elem. Sch. J.*, 1930, **31**, 183–190.
11. Peterson, J. C. The higher mental processes in learning. *Psychol. Monogr.*, 1920, **28**, No. 7 (Whole No. 129).
12. Ragsdale, C. E. How children learn the motor types of activities. In N. B. Henry (Ed.) *Yearb. nat. Soc. Stud. Educ.*, 49, Part I. *Learning and Instruction.* Chicago: Univer. of Chicago Press, 1950.
13. Ruger, H. A. The psychology of efficiency. *Arch. Psychol.*, 1910, **2**, No. 15.
14. Thiele, C. L. The contribution of generalization to the learning of addition facts. *Teach. Coll. Contr. Educ.*, 1938, No. 763.
15. Waters, R. H. The influence of tuition upon ideational learning. *Psychol. Monogr.*, 1920, **28**, No. 7 (Whole No. 129).

Received February 6, 1957.

STUDENT VS. INSTRUCTOR DESIGN OF STUDY METHOD[1, 2]

SLATER E. NEWMAN[3]

Air Force Personnel and Training Research Center, Colorado

The assigning of study tasks to be done by the student either in or out of class is common educational practice. Such practice, however, implies an assumption that each student can best prescribe for himself how such material should be studied.

It was the purpose of the present investigation to demonstrate that such an assumption is not warranted; that rather, the use of study materials and procedures designed on the basis of information provided by research on learning will lead to faster student learning than the use by each student of his own set of techniques.

Procedure

The task confronting each student was that of learning the names for each of 20 electrical symbols. This task was chosen because it appeared to be representative of the kind of task on which considerable research-based information already exists. The following schedule was used with each *S*, all of whom were run individually:

1. A pretest was administered to determine how many of the 20 symbols the student could name correctly. The papers of those students who knew the names of more than three symbols were not used in the analysis of posttest scores.

---

[1] This report is based on work done under ARDC Project No. 7715, Task No. 77283, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government. The opinions or conclusions expressed herein are those of the author. They are not to be construed as necessarily reflecting the views or endorsement of the Department of the Air Force or of the Air Research and Development Command.

[2] A paper, based on some of the findings reported here, was presented at the 1956 meetings of the American Psychological Association.

[3] Now with the Department of Psychology, North Carolina State College, Raleigh, North Carolina.

UNIVERSITY OF MICHIGAN LIBRARIES

2. Instructions dealing with what the student should do during the learning period were read aloud by the instructor.

3. Each student studied the list either for 7¾ or for 12 minutes. It was anticipated that such differences in study time would produce statistically significant differences in performance on the posttest and that these differences might be used as a rough yardstick in gauging the effects of the kind of procedure used during the study session.

4. The posttest (which was the same as the pretest) was then administered. All 20 symbols were presented on a sheet of paper, and the student was asked to write the name of each symbol he knew to the right of the symbol. The posttest was administered immediately after the study period.

Students were randomly assigned either to a Student Group or to an Instructor Group. The procedures used during the study period with each group were the following:

Student Group ($N = 15$). At the beginning of the study period, the student was given a sheet on which were listed each of the 20 electrical symbols and its name, and he was asked to study the list. Each name was to the right of its symbol. The order in which these symbols were arranged on the study sheet differed from that in which they appeared on the test sheet. Seven students were allowed 7¾ minutes for study, and eight students were allowed 12 minutes for study. The instructor was present during the entire study period.

Instructor Group ($N = 15$). Suggestions for techniques to be used with this group were obtained from five psychologists familiar with the results of research on learning the kind of task employed here. To each psychologist the symbol-name pairs to be learned were presented, and he was asked to suggest techniques which could be expected to facilitate the learning of these pairs. There was not, however, complete agreement among the five psychologists as to techniques which should be employed. Thus, the combination of techniques used with the Instructor Group was based on a selection by the author from among those suggested. For the final procedure, file cards were prepared on which appeared either the symbol and its name, the symbol alone, the name alone, or the name of the set (to be described next) to which a symbol belonged. The 20 symbols were divided into four sets of five each. All symbols in a set shared one feature in common. For set one, it was a circle; for set two, a wavy line; for set three, a black dot; and for set four, a straight line. Early in training, study of set one was completed before study of set two was begun, etc. Approximately 10 seconds

STUDENT VS. INSTRUCTOR DESIGN OF STUDY METHOD[1, 2]

SLATER E. NEWMAN[3]

Air Force Personnel and Training Research Center, Colorado

The assigning of study tasks to be done by the student either in or out of class is common educational practice. Such practice, however, implies an assumption that each student can best prescribe for himself how such material should be studied.

It was the purpose of the present investigation to demonstrate that such an assumption is not warranted; that rather, the use of study materials and procedures designed on the basis of information provided by research on learning will lead to faster student learning than the use by each student of his own set of techniques.

Procedure

The task confronting each student was that of learning the names for each of 20 electrical symbols. This task was chosen because it appeared to be representative of the kind of task on which considerable research-based information already exists. The following schedule was used with each *S*, all of whom were run individually:

1. A pretest was administered to determine how many of the 20 symbols the student could name correctly. The papers of those students who knew the names of more than three symbols were not used in the analysis of posttest scores.

---

[1] This report is based on work done under ARDC Project No. 7715, Task No. 77283, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government. The opinions or conclusions expressed herein are those of the author. They are not to be construed as necessarily reflecting the views or endorsement of the Department of the Air Force or of the Air Research and Development Command.

[2] A paper, based on some of the findings reported here, was presented at the 1956 meetings of the American Psychological Association.

[3] Now with the Department of Psychology, North Carolina State College, Raleigh, North Carolina.

UNIVERSITY OF MICHIGAN LIBRARIES

2. Instructions dealing with what the student should do during the learning period were read aloud by the instructor.

3. Each student studied the list either for 7¾ or for 12 minutes. It was anticipated that such differences in study time would produce statistically significant differences in performance on the posttest and that these differences might be used as a rough yardstick in gauging the effects of the kind of procedure used during the study session.

4. The posttest (which was the same as the pretest) was then administered. All 20 symbols were presented on a sheet of paper, and the student was asked to write the name of each symbol he knew to the right of the symbol. The posttest was administered immediately after the study period.

Students were randomly assigned either to a Student Group or to an Instructor Group. The procedures used during the study period with each group were the following:

Student Group ($N = 15$). At the beginning of the study period, the student was given a sheet on which were listed each of the 20 electrical symbols and its name, and he was asked to study the list. Each name was to the right of its symbol. The order in which these symbols were arranged on the study sheet differed from that in which they appeared on the test sheet. Seven students were allowed 7¾ minutes for study, and eight students were allowed 12 minutes for study. The instructor was present during the entire study period.

Instructor Group ($N = 15$). Suggestions for techniques to be used with this group were obtained from five psychologists familiar with the results of research on learning the kind of task employed here. To each psychologist the symbol-name pairs to be learned were presented, and he was asked to suggest techniques which could be expected to facilitate the learning of these pairs. There was not, however, complete agreement among the five psychologists as to techniques which should be employed. Thus, the combination of techniques used with the Instructor Group was based on a selection by the author from among those suggested. For the final procedure, file cards were prepared on which appeared either the symbol and its name, the symbol alone, the name alone, or the name of the set (to be described next) to which a symbol belonged. The 20 symbols were divided into four sets of five each. All symbols in a set shared one feature in common. For set one, it was a circle; for set two, a wavy line; for set three, a black dot; and for set four, a straight line. Early in training, study of set one was completed before study of set two was begun, etc. Approximately 10 seconds

intervened between study of each set. These materials were used in a five-step procedure as follows:

Step 1. Each of the five names was presented individually for approximately five seconds (e.g., capacitor). The instructor pronounced each name and then had the student pronounce the name.

Step 2. The set name was then presented (e.g., straight lines). For this and subsequent steps, a three-second rate of presentation was employed for each card.

Step 3. Each of the five symbol-name cards was then presented (e.g., capacitor symbol-capacitor). The student pronounced each name and was corrected if he erred.

Step 4. Each card with the symbol alone was then presented (e.g., capacitor symbol), and the student responded with the name of the symbol if he felt he knew the name.

Step 5. After these steps had been followed with each of the four sets, each of the 20 symbols was presented alone and was followed by a card on which appeared the symbol and its name.

With seven students, this set of procedures was used once (total study time for each student approximately 7¾ minutes). With eight additional students, this set of procedures, excluding Step 1 (i.e., response learning) was then repeated a second time. The total study time for this group was approximately 12 minutes.

Subjects

Thirty airmen, each of whom had recently completed basic training, were randomly assigned to one of the experimental groups. All *S*s had a score of 5 or above on the Airman Classification Battery's Technical Specialty Aptitude Index (a measure of general intellectual ability).

Results

The random assignment of students to groups did not result in significant between-group differences in aptitude. Nevertheless, each student's score on the posttest was adjusted on the basis of his aptitude index. A 2 × 2 analysis of variance (i.e., type of study × amount of study time) of these adjusted scores was then accomplished and is presented in Table 1. The *F*s for both type of study and amount of study were found to be significant at the .01 level. The mean scores for each cell of the design are presented in Table 2.

The significant *F* for type of study indicates that there were

UNIVERSITY OF MICHIGAN LIBRARIES

TABLE 1

ANALYSIS OF VARIANCE OF ADJUSTED SCORES

Source	*df*	Mean Square	*F*	*p*
Type of study	1	10,046.70	10.51	.01
Amount of study time	1	12,080.38	12.64	.01
Interaction	1	234.00	.24	
Within	26	955.50		
Total	29			

TABLE 2

MEAN APTITUDE AND RETENTION TEST SCORES FOR EACH GROUP

Group	Amt. of Study Time	*N*	Mean Aptitude Index	Mean Retention Score (Unadjusted)	*SD*[a]
Instructor	7¾ min.	7	7.4	9.6	5.33
Student	7¾ min.	7	6.7	13.1	1.59
Instructor	12 min.	8	6.8	13.5	2.91
Student	12 min.	8	7.1	17.0	2.00

[a] Bartlett's chi square for these data is 7.36 ($.10 > p > .05$).

real differences between the average performance scores of the Student and the Instructor Groups. The differences, however, were not in the direction anticipated. Instead, students who studied using their own techniques did better on the posttest than students for whom study techniques were prescribed.

A second study was then conducted to determine whether the introduction of certain changes in the procedure used with the Instructor Group would result in faster learning by *S*s of this group than by *S*s who studied using their own set of techniques. Two changes in procedure were introduced with the Instructor Group: (*a*) prompting training was used throughout (i.e., immediately following presentation of a stimulus-response pair, the stimulus-term was presented alone, *S* being required to respond with its response-term); thus, *S* was never tested during training, and (*b*) the items were presented in a fixed order, both within sets and between sets. This revised set of procedures was used with 26 *S*s. A Student Group ($N = 26$) received the same treatment as the Student Group in the first study. The *S*s in each group were

allowed 7¾ minutes for study. A median test applied to the posttest scores of both groups resulted in a chi square of 7.69 ($p < .01$), the performance of the Student Group again exceeding that of the Instructor Group.

Discussion

An examination of the differences in procedures used with the Student and the Instructor Groups might shed some light on these performance differences. Of particular interest is the fact that for members of the Student Groups, all symbols and names were present throughout the study period. The implications here are that (*a*) this probably allowed for more opportunity to observe differences between symbols and between names than was the case with the Instructor Groups; and (*b*) students could spend more time studying those symbols and names with which they were having difficulty.

A second difference between the groups is that members of the Student Groups were probably more familiar with the techniques which were used (and which each student selected) during the learning period than were students of the Instructor Groups. Thus, the posttest performance differences between the Student and Instructor groups may not have been so great had students of the Instructor Groups had some practice in learning how to learn from the techniques used during their study period.

Finally, the manner in which the techniques were combined for use with the Instructor Groups may have led to slower learning (e.g., the relatively little emphasis on response learning).

At least two conclusions *might* appear warranted on the basis of the results of these studies: (*a*) that allowing the student to decide on the techniques he will use during study on tasks like that of learning the names for electrical symbols is good educational practice, and (*b*) that in respect to the design of materials and procedures for learning on such tasks, the experimental psychology of learning has little to offer. Such conclusions do not appear to present a complete picture.

The situation, rather, may be described like this: The effective design of materials and procedures for tasks even so apparently simple as learning the names of electrical symbols is not easy. Much information is already available from laboratory research on learning which might prove useful in designing materials for

UNIVERSITY OF MICHIGAN LIBRARIES

such tasks. However, until considerable experience is accumulated about how such information can best be applied to the design of materials and procedures for learning in classroom type situations (or until additional research-based information is available), the practice of allowing the student to decide on the techniques he will employ during study of the kind of task used here may be a defensible one.

Received December 10, 1956.

RETROACTIVE INHIBITION AND FACILITATION IN THE LEARNING OF SCHOOL MATERIALS

DAVID P. AUSUBEL, LILLIAN CUKIER ROBBINS, AND ELIAS BLAKE, JR.

University of Illinois

This study is generally concerned with some of the major determining factors and processes involved in the retention of typical school materials. More specifically, it deals with (*a*) the relative importance of proactive and retroactive inhibition in the forgetting of such materials, (*b*) the role of similarity between original and interpolated tasks in retroactive inhibition, and (*c*) the relative facilitating effects of simple repetition and explicit retroactive enhancement of the discriminability of learning material on its subsequent retention.

The paradigm of retroactive inhibition is congruent with the highly paradoxical "interference" theory of forgetting. According to this theory, the introduction of *identical* material subsequent to learning results in consolidation and an increment in performance but the more *similar* the subsequently introduced material is to the original learning task, short of complete identity, the more it supposedly interferes with the retention of the latter. The resulting decrement in performance (as well as the interfering processes responsible for the decrement) is designated as "retroactive inhibition."

Practically all of the experimental work on retroactive inhibition has been conducted with materials (e.g., nonsense syllables, lists of paired adjectives) that are quite far removed from the type of learning that occurs in the classroom. This approach has not been completely without rationale. First, nonsense syllables, as additive units of equal difficulty, make for greater quantifiability. Second, because they are relatively meaningless, their use makes it possible to rule out the indeterminable influence of meaningful antecedent conditions that obviously vary greatly from one individual to another. Unfortunately, however, although the use of nonsense syllables adds undoubted methodological rigor to the study of retroactive inhibition, their very nature limits the applicability of findings to the retention of meaningless materials—or to a type of

UNIVERSITY OF MICHIGAN LIBRARIES

learning that is rare both in everyday situations and in the school room. Nevertheless, despite the absense of any convincing a priori grounds for believing that retroactive inhibition occurs in the same way for meaningful as for relatively meaningless (arbitrary, rote, discrete) learning, findings from the latter kind of situation have been commonly and uncritically extrapolated to the former.

What reasons are there for suspecting that retroactive inhibition might operate differently for meaningful than for meaningless materials? First, nonsense syllables are discrete and isolated entities that are almost completely unrelated to the vast reservoir of concepts and information that a pupil brings with him to a learning situation. They are relatively unaffected by previous learnings; and, hence, any forgetting that results from the studying of subsequently introduced materials (over and above the forgetting of an equated group that "rests" between initial test and retest) can be plausibly attributed to the interfering influence of these latter materials. This is simply not the case in school learning. Almost any new meaningful material is related in some way to what has been learned in the past; in fact, the curriculum is deliberately organized in this fashion to provide for the gradual and untraumatic introduction of new facts and concepts. Hence, it seems reasonable to suppose that any school learning will involve interaction with established knowledge in the learner, and that this interaction will probably have greater influence on retention than the immediate interactive effects of material introduced subsequent to such learning. Furthermore, the retroactively introduced materials also have significant connections with the residue of prior learning experience. Thus, whatever effect they do have on retention is probably less a reflection of interaction with the learning task in their own right than of the activation and mobilization of relevant subsuming concepts in the individual's existing cognitive structure.

Second, in dealing with meaningless material that cannot be related to established concepts in cognitive organization, a simple paradigm of retroactive interference, e.g., stimulus or response generalization, can adequately explain forgetting. Meaningful materials, however, are invariably related to an existing cognitive structure that is hierarchically organized in terms of highly stable and inclusive conceptual clusters under which are subsumed less stable and more specific illustrative data. As they enter the cognitive field, they are appropriately "catalogued" under relevant

conceptual foci, since this is the most orderly and efficient way of organizing large quantities of material for ready availability. At first, for a variable period of time, the recently catalogued data can be dissociated from their subsuming concepts and are reproducible as individually identifiable entities. Because it is most economical and least burdensome, however, to subsume as much material under as inclusive categories as possible, the import of many specific illustrative items is assimilated by the generalized meaning of the more established and highly conceptualized subsuming foci. When this happens, these items are no longer dissociable or identifiable as entities in their own right and are said to be "forgotten." Hence, barring repetition or some other special reason, e.g., primacy, uniqueness, affective content, for the perpetuation of dissociability strength, specific items of experience that are supportive of an existing conceptual entity tend eventually to be irreversibly subsumed by the generality of the latter. In other words, it is proposed that subsumability under established concepts, rather than retroactive interference due to similarity per se, is the determining factor in the forgetting of meaningful learning material.

Third, similarity between original and interpolated materials presumably bears a different relationship to retroactive inhibition when the materials involved consist of conceptual data and related information than when they consist of rote, nonsense tasks. In the latter instance, specific differences in content, in the absence of any transferable conceptual meaning, completely obviate the possibility of retroactive facilitation and only generate confusion. It is little wonder, then, that degree of similarity (short of identity) between original learning task and interpolated material is positively related to degree of retroactive inhibition. Meaningful materials, however, in addition to being characterized by specific detail and context, have a more general conceptual content that is transferable to comparable learning situations. Hence, contextually dissimilar but substantively similar interpolated learning can retroactively facilitate the retention of previously learned materials.

A final theoretical issue that we shall explore in this study has to do with the mechanism of retroactive facilitation in the retention of meaningful material. By retroactive facilitation is meant the effect of interpolated material in *reducing* the retention loss caused by proactive inhibition. It is demonstrated when, after interpolated

UNIVERSITY OF MICHIGAN LIBRARIES

learning, significantly less loss occurs from immediate test of original retention to later retest than occurs when no learning task is interpolated between test and retest.[1] Consideration of the "subsumption-dissociability" theory of forgetting proposed above suggests that a given ideational element is irreversibly subsumed under an established concept not because of similarity between them per se, but because its distinguishing features are not sufficiently salient to enable it to persist as a discriminably distinct entity in its own right. Thus, it is conceivable that identical or substantive repetition of the learning material induces retroactive facilitation by increasing its discriminability from already established concepts. Another possible way of retroactively enhancing the discriminability of learning material is explicitly to point out similarities and differences between it and potentially subsuming concepts. Since this latter method presumably results in more explicit appreciation of the distinctive properties of the learning material, it should lead to greater retroactive facilitation than an equivalent amount of simple repetition.

Hypotheses

Proceeding from the foregoing theoretical considerations, therefore, an experiment was designed to put to empirical test the following hypotheses regarding the retention of meaningful school materials: (*a*) Proactive inhibition, i.e., irreversible subsumption of learning material by established concepts, accounts for more forgetting than does retroactive inhibition. (*b*) The interpolation of substantively similar (but not identical) material between initial test and retest results in retroactive facilitation rather than retroactive inhibition. (*c*) Less retroactive facilitation occurs when the interpolated material merely recapitulates the learning passage than when it explicitly delineates similarities and differences between the latter and presumed subsuming concepts.

Method

Subjects

An experimental group and three control groups were constituted out of a total population of 188 undergraduates in eight sections

[1] If sufficient retroactive facilitation occurs between test and retest, as when the learning task is repeated frequently during the interpolated sessions, an actual increment rather than a reduced decrement results.

of a course in educational psychology at the University of Illinois. The mean ages of these groups varied from 21.8 to 22.4. The experiment was rationalized as a laboratory exercise in the psychology of learning and was performed during regularly scheduled class hours. In order to maximize ego involvement, the *S*s were informed that after the data were processed their individual scores, as well as the class results, would be reported to them.

Learning Material and Test of Retention

The learning material used in this study was a passage of approximately 1700 words on the history, sacred literature, doctrine, and ethical teachings of Buddhism. It was specially designed for use in an experimental learning situation that would bring into play the same kinds of cognitive variables and interactional processes that affect the long-term retention of typical ideational material learned in a school setting. Hence, it had to satisfy the major criterion of being largely unfamiliar in its specific aspects but still generally related to a body of knowledge relatively well established in most of our *S*s. On purely a priori grounds it seemed reasonable to assume that the facts and ideas contained in this passage would interact with and be subsumed under existing concepts of Christianity, and that the latter concepts would furnish the basis for whatever proactive and retroactive inhibition might occur in the course of learning and retention.

A 37-item multiple-choice test was constructed to test immediate and later retention of the Buddhism material. Preliminary tryouts of the passage and the test indicated that the material was actually unfamiliar but learnable, and that the test was reliable and structurally sound. Scores on the test showed a satisfactory range of variability and were distributed normally. The corrected odd-even coefficient of equivalence was .81. When the test was administered to naive *S*s who had not previously studied the passage, the mean score (number of correct items) was only 8.9 as compared to a chance expectancy score of 7.4 (one-fifth of 37).

Because of the limited duration of class periods, the passage could not be made long enough to provide a sufficient number of conceptual items for two reasonably adequate tests. Practical considerations, therefore, made it necessary to use the same test as a measure of both immediate and later retention. Hence, it is conceivable that memory of answers given initially competed in some

UNIVERSITY OF MICHIGAN LIBRARIES

instances on the retest with the *Ss*' recollections of the material itself. The probability of this happening, however, was minimized by the eight-day interval between tests and the difficulty of specifically remembering one of five highly similar and complexly worded alternatives. In any event, this factor operated in similar ways for the experimental and control groups.

Experimental Procedure

The *Ss* in all groups were first instructed to read the Buddhism passage at their normal reading speed (eight to ten minutes) and then to utilize the remainder of the 25 available minutes to commit the concepts and facts in the passage to memory. It was explained that a multiple-choice test would be given immediately after the reading was completed and also eight days later. Both tests were given as announced.

Twenty-four hours after the initial test, 20-minute sessions of interpolated learning were introduced as follows: The Experimental Group studied a 2100 word comparative essay on Buddhism and Christianity. This essay encompassed all of the material included in the Buddhism passage, but similarities and differences between the two religions were explicitly pointed out. Control Group A restudied the Buddhism passage. Control Group B studied a 1700 word passage on the history, sacred literature, doctrine, and ethical teachings of Christianity. In each instance it was explained that knowledge of the material studied during the interpolated session would be tested for a week later by the previously announced retest. Control Group C was not subjected to an interpolated learning session.

Results and Interpretation

Means and standard deviations of all groups on the initial test and retest are shown in Table 1. Since raw scores on this multiple-choice test include correct answers attributable to guessing, incidental knowledge, and the presence of unintentional cues in the five alternative answers, corrected means were computed by subtracting from the raw means the mean score of *Ss* who took the test without studying the Buddhism passage. These corrected means were utilized in making all intragroup and intergroup comparisons of test scores. The distribution of test scores did not diverge significantly from the normal. F tests for homogeneity of

TABLE 1

MEANS AND STANDARD DEVIATIONS OF IMMEDIATE AND LATER TESTS OF RETENTION FOR EXPERIMENTAL AND CONTROL GROUPS

Group	Interpolated Material	*N*	Test I			Test II			Percentage Retained[b] (Test II/ Test I)
			Raw Mean	Corr. Mean[a]	*SD*	Raw Mean	Corr.[a] Mean	*SD*	
Experimental	Comparative	43	24.4	15.5	5.47	25.6	16.7	5.25	107.7
Control A	Identical	50	22.5	13.6	6.14	23.2	14.3	5.75	105.1
Control B	Subsuming	44	24.7	15.8	4.79	21.9	13.0	4.83	82.3
Control C	None	51	24.5	15.6	5.84	22.0	13.1	6.13	84.0

[a]Corrected means were computed by subtracting 8.9 (the mean score of *Ss* who took the test without studying the learning material) from the raw means.
[b]Corrected means were utilized in computing the percentage retained from test to retest.

variance between groups indicated that we were dealing with comparable groups (i.e., drawn from the same population).

The relatively high degree of retention (84%) demonstrated by Control Group C after an eight-day interval without any intervening repetition can be attributed to several factors. First, the gradient of forgetting is known to be very gradual when meaningful material is memorized. Second, the availability of memory traces is maximized when a recognition rather than a recall method of retention is employed. Third, a type of reminiscence effect was partially operative as frequently becomes evident when immediate retention of partially learned, highly conceptualized material is compared to retention several days later. Dissociability of newly subsumed materials presumably tends to be at a minimal level immediately after the initial shock of interaction with their subsuming concepts, and improves spontaneously for a limited time thereafter as the temporary confusion generated by the interaction gradually dissipates. Thus, although the eight-day interval in this instance undoubtedly exceeded the temporal limits of the spontaneous recovery phase, the ratio of later to immediate retention was probably still inflated by the temporary inhibition of retention immediately after the completion of learning.

Proactive versus Retroactive Inhibition

To test the hypothesis that most forgetting of meaningful material is reflective of proactive rather than retroactive inhibition, the retest scores and percentages of retention (Test II/Test I) of control groups B and C were compared (see Table 1). Control Group B had studied an interpolated passage on Christianity, i.e.,

UNIVERSITY OF MICHIGAN LIBRARIES

the material potentially most capable of inducing retroactive inhibition in relation to the Buddhism learning passage. Control Group C had not been exposed to an interpolated learning session. Thus, in the first instance, both retroactive and proactive inhibition were operative, whereas in the second instance only proactive inhibition was operative. The data show conclusively that subsequent study of relevant subsuming material results in only negligibly more forgetting of the original learning material than when the latter material interacts spontaneously with established subsuming concepts that the individual brings with him to the learning situation. Since most ideational material that is learned in a school setting is related to and hence potentially subsumable by an existing body of stable concepts in the learner's cognitive organization, this finding is generally applicable to school learning. It suggests that retroactive inhibition is an important determining factor in forgetting only in artificial laboratory situations employing nonsense learning tasks, and is at most a relatively minor variable in the retention of most newly learned school materials.

The Role of Similarity in Retroactive Inhibition

That similarity short of identity between learning and interpolated tasks does not cause retroactive inhibition in the case of meaningful material, but may or may not induce retroactive facilitation, is evident from intergroup comparisons of the retention data of the Experimental Group and Control Groups B and C (see Table 1). This type of analysis supports the view that the influence of interpolated learning on retention is a function of whether or not it increases the discriminability of the original learning material and thereby counteracts irreversible subsumption of the latter by established concepts. As already pointed out, although interpolation of similar subsuming material (Christianity passage) apparently does not induce retroactive inhibition, neither does it reduce proactive inhibition. A t test between the retest scores (or the percentages of retention) of the Experimental Group and Control Group C, on the other hand, shows that the interpolation of similar comparative material results in a highly significant amount of retroactive facilitation ($p < .001$). Despite the seven-day interval between interpolation and retest, slightly more (7.7%) material was remembered on Test II than on Test I. The net conceptual residue of the comparative passage, which consisted

of recapitulated Buddhism material, subsuming Christianity material, and an explicit comparison between the two religions, unequivocally increased the discriminability of the original Buddhism material. It appears, therefore, that in dealing with meaningful school learning, the absence of point-for-point-identity between original and interpolated learning tasks not only fails to induce retroactive inhibition but is also compatible with the occurrence of retroactive facilitation as long as the interpolated material enhances the discriminability of the learning task.

Repitition versus Explicit Comparison

To what can the facilitating effects of the interpolated comparative passage be attributed? One possible explanation, emphasized in our theoretical discussion above, is that explicit delineation of the similarities and differences between the learning passage (Buddhism) and its presumed subsuming concepts (Christianity) enhanced the discriminability of the former and made it more resistive to irreversible subsumption by the latter. However, our experimental design had to take into account the obvious fact that a comparative presentation must of necessity recapitulate in a different context the very same ground covered by the original learning material. It is patently impossible to compare Buddhism to Christianity without discussing Buddhism. Hence, it could be plausibly argued that the retroactive facilitation following interpolation of the comparative passage was largely induced by the recapitulative features of the latter material; and since there was, in fact, no significant difference between the Experimental Group (comparative interpolation) and Control Group A (repetitive interpolation) in the ratio of later to immediate retention scores (see Table 1), this argument appears incontrovertible. It must be concluded, therefore, (*a*) that most if not all of the retroactive facilitation associated with the interpolation of the comparative passage can be attributed to substantive repetition of the learning passage in another context, and (*b*) that *retroactive* use of comparative material is no more efficacious in counteracting proactive inhibition than simple recapitulation of the learning material itself. These conclusions, of course, do not detract from the finding that, where meaningful learning is involved, interpolation of nonidentical material not only does not necessarily result in retroactive inhibition but may also induce retroactive facilitation.

UNIVERSITY OF MICHIGAN LIBRARIES

The failure of the comparative approach to reduce proactive inhibition any more than was accomplished by simple repetition can be attributed perhaps to the fact that this method may only be effective in *preventing* irreversible subsumption caused by *initial* lack of discriminability. Once material, by virtue of such initial limitations in discriminability, can no longer be dissociated from its subsuming concepts at the time of original learning, repetition may be the only possible way of retroactively enhancing discriminability. This explanation is being tested presently by means of a research design which contrasts the retention of *S*s initially learning the comparative passage with the retention of matched *S*s initially learning the Buddhism passage.

Summary and Conclusions

Retroactive inhibition and facilitation in the retention of meaningful school materials was studied in a population of 188 undergraduates divided into an experimental and three control groups. All *S*s studied a 1700-word passage on Buddhism and were tested by means of a multiple-choice examination for retention of same immediately after learning and eight days later. Interpolated learning sessions were instituted 24 hours after original learning and were arranged as follows: The Experimental Group learned a passage explicitly comparing Buddhism and Christianity. One control group restudied the Buddhism passage, and another control group studied a passage dealing solely with Christianity. The remaining control group was not subjected to any interpolated learning.

Intergroup comparisons of the ratios between initial and later retention supported the hypotheses that, where meaningful school material is involved, (*a*) proactive rather than retroactive inhibition is the determining factor in forgetting; and (*b*) interpolated learning of material substantively similar to but not identical with the original learning task does not result in retroactive inhibition, but induces as much retroactive facilitation as identical repetition of the learning material. Explicit comparison of the learning material and its presumed subsuming concepts was no more effective than simple repetition of the learning material in inducing retroactive facilitation. The findings were interpreted in the light of a "subsumption-dissociability" theory of retention and forgetting.

Received December 17, 1956.

THE IDENTIFICATION OF NUMBER ANXIETY IN A COLLEGE POPULATION

RALPH MASON DREGER

Jacksonville University

LEWIS R. AIKEN, JR.

Emory University

Many persons report in clinical sessions and in academic classes that they are emotionally disturbed in the presence of mathematics. It has been the writers' experience that reactions of this nature to mathematical and arithmetical materials constitute a major field of emotionality in academic situations. And yet, although many studies have been conducted to determine optimum ways of presenting arithmetic materials (**7**), almost no controlled research has been attempted in the realm of emotional problems associated with arithmetic and mathematics. Schonell (**12**) observed that backwardness in arithmetic is due as much to emotional as to intellectual factors, and psychoanalytic writers (**6, 8, 11**) suggest that failure in arithmetic may be related to maternal overprotection.

Inasmuch as there has not been any attempt to determine if actually there is a set of reactions to numbers and mathematics on the scale that appeared evident to the writers, an inquiry was called for to investigate the extent and nature of such reactions. Consequently, the study reported here is an endeavor to detect the presence of a syndrome of emotional reactions to arithmetic and mathematics, tentatively designated "Number Anxiety." We speak of a syndrome, for anxiety is probably more accurately described as "anxieties" rather than "anxiety" (**2**). The existence of specific situational anxieties has been shown in several investigations (**3, 13**). Moreover, O'Connor's (**10**) finding of five distinct factors on the Taylor Scale of Manifest Anxiety shows that several dimensions of anxiety are measured by this popular scale. Number Anxiety, too, may turn out to be a complex variable.

Hypotheses

Three hypotheses were formulated on the basis of previous observation concerning the existence of Number Anxiety and its relations to concomitant human functioning:

UNIVERSITY OF MICHIGAN LIBRARIES

1. Number Anxiety exists as a factor distinct from (although it may accompany) "general anxiety."
2. Number Anxiety is not related to intelligence as measured by the Wechsler-Bellevue or ACE scales, although it is probably negatively correlated with ACE Q scores.
3. Persons with high Number Anxiety will tend to make lower scores in mathematics than others, even though intelligence (as measured above) between these two groups is not significantly different.

Subjects

Seven hundred and four students in basic mathematics classes at Florida State University.

Procedure

The Taylor Manifest Anxiety Scale (**14**), omitting three items of relatively low validity and including three items (Table 1) specifically designed to measure feelings of anxiety concerning working with numbers, was administered to the 704 students at the close of class periods during which mathematics tests had been returned to them. The *S*s were informed that the results of the Taylor Scale, called the "Confidential Data Sheet," would be used for scientific purposes and would have no influence on their grades or careers.

Scores on the three Number Anxiety items were correlated separately and together with the total score on the 47-item Taylor scale. In addition, intercorrelations were obtained among the three Number Anxiety items and the 14 items on the Taylor scale shown in two previous studies (**1, 5**) to be the most valid, making a 17 × 17 correlation matrix. This matrix was subjected to a Holzinger and Harman cluster analysis (**4**).

According to total scores on the three-item Number Anxiety items and the 47-item Taylor Scale, the 704 *S*s were divided into four groups, those high in both "general anxiety" and Number Anxiety, those high in one and low on the other and vice versa, and those low in both. From each of these four groups ten *S*s were chosen by a table of random numbers for further investigation. Inasmuch as males and females were not significantly different in scores on either scale, the sex variable was randomized. The 40 students in the subsample were not informed of the basis of

selection, and were presumably ignorant of the connection between the administration of the "Confidential Data Sheet" and the following procedures. The *Ss* in the subsample were given Form I of the Wechsler-Bellevue Intelligence Scale (**15**) by one examiner. At the same time as the administration of the verbal subtests, pre- and poststimulus readings were taken on a psychogalvanometer registering palmar skin resistance (**9**). Besides the recording of GSR deflections for each subtest item, the deflection was noted for the arithmetic subtest instructions, "I want to see how good you are in arithmetic."

The final Mathematics 105 grades for the 704 students and their American Council on Education Psychological Examination scores, Q, L, and Total, were obtained and correlated with Number Anxiety scores. Correlations were also found comparing grades, ACE, and Wechsler-Bellevue Verbal, Performance, and Full Scale IQ's with Number Anxiety scores for the 40 *Ss* investigated more intensively.

Results

Number Anxiety as defined by the three scale items (Table 1) proved to be present in approximately one-third of the 704 students; 239 or 35% responded positively to two or three of the key items.

With some minor differences, scores on the 47-item Taylor scale formed a distribution very similar to that of Taylor's original score distribution. Like the original distribution, the one obtained in this study is positively skewed; it was assumed, however, that it probably represents a normal distribution sufficiently well to allow correlations which employ a dichotomy of scores. Correla-

TABLE 1

BISERIAL CORRELATIONS BETWEEN THREE "NUMBER ANXIETY" ITEMS AND TOTAL SCORE ON TAYLOR SCALE SPLIT AT MEAN

Item	Correlation with Taylor Scale
3. I am often nervous when I have to do arithmetic.	.32
9. Many times when I see a math problem I just "freeze up."	.37
38. I was never as good in math as in other subjects.	.09
Total, items 3, 9, 38 combined	.33

UNIVERSITY OF MICHIGAN LIBRARIES

TABLE 2
FACTOR LOADINGS IN "LITTLE FACTOR ANALYSIS" OF THREE "NUMBER ANXIETY" ITEMS

Item	Factors	
	A	B
3	−.002	.552
9	.266	.324
38	.518	.001

tions between the Taylor scale and the Number Anxiety items, individually and *in toto*, are presented in Table 1.

The relatively high correlations among the Number Anxiety items (Table 3) led the writers to suppose that they represent a cluster or syndrome; but the fact that they are not any more highly intercorrelated further suggested that the three items are not factorially pure. With the full recognition that a factor analysis of three items is purely an exploratory measure, a factor analysis, which was termed a "little factor analysis" to distinguish it from a serious one, was performed employing only the Number Anxiety items 3, 9, and 38, by means of Thurstone's centroid method. Two factors accounting for all of the common factor variance emerged. Orthogonal rotation allowed for no psychological meaningfulness, but oblique rotation to simple structure (as far as the latter can be obtained with three items) yielded greater significance (see Table 2). Factor A was identified very tentatively as "Negative math. reaction" and Factor B as "Nervousness in the presence of math."

A cluster analysis (**4**) of the correlation matrix (Table 3) of 14 Taylor Scale and three Number Anxiety items revealed three distinct clusters, tentatively labelled: I. Compulsive inattentiveness (composed of items 6 and 41 from the Taylor Scale); II. Covert, conscious strain (items 47, 45, 42, 48, 40, 44, 21, 5, 27, 29, 49, and 34); and III. Nervousness and difficulty in arithmetic (items 3, 9, and 38). The second cluster approaches "general anxiety" and the third clearly represents what is called here Number Anxiety.

Log conductance changes of the GSR (**9**) subjected to analysis of variance revealed that only on the arithmetic subtest and arithmetic instructions were there significant *F* ratios (at the .05

TABLE 3
TETRACHORIC CORRELATIONS AMONG 17 ITEMS FROM CONFIDENTIAL DATA SHEET

	5	6	9	21	27	29	34	38	40	41	42	44	45	47	48	49
3	.41	.10	.67	.37	.33	.24	.12	.50	.20	.03	.14	.28	.32	.35	.31	.10
5		.44	.37	.70	.59	.29	.37	.09	.52	.30	.42	.52	.67	.50	.54	.23
6			.14	.26	.28	.32	.31	.02	.40	.82	.40	.31	.50	.51	.38	.31
9				.33	.36	.16	.22	.64	.21	.18	.17	.06	.24	.19	.24	.12
21					.37	.29	.53	.27	.13	.31	.40	.50	.71	.34	.60	.27
27						.43	.38	.59	.45	.29	.56	.59	.62	.55	.55	.33
29							.16	.08	.30	.26	.32	.47	.47	.38	.28	.28
34								−.09	.46	.37	.36	.41	.31	.39	.29	.10
38									−.01	−.08	−.03	.01	−.02	.14	−.01	.07
40										.48	.47	.35	.48	.57	.68	.39
41											.48	.30	.48	.52	.47	.36
42												.47	.50	.58	.45	.38
44													.58	.37	.51	.13
45														.58	.52	.37
47															.55	.55
48																.49

level). That is, on Information, Comprehension, Digit Span, and Similarities there were no significant differences between Number Anxiety and "general anxiety" individuals or between high and low scorers on each anxiety scale. None of the interactions was significant, even on the two arithmetic comparisons. A breakdown of the arithmetic and arithmetic instructions GSR changes showed that all of the significant differences in each case lay between high and low scorers on Number Anxiety.

Correlations between Number Anxiety and final grades in Mathematics 105 for both the 40 subsample group and the 704

TABLE 4
BISERIAL CORRELATIONS BETWEEN "NUMBER ANXIETY" AND FINAL GRADES IN MATH, WECHSLER-BELLEVUE IQS, AND ACE SCORES

No. of Subjects	Grades	Wechsler-Bellevue			ACE		
		PS	VS	FS	Q	L	T
40	−.51	.20	−.015	.44	−.25	−.08	−.04
704	−.44						

UNIVERSITY OF MICHIGAN LIBRARIES

original sample, together with the correlations of Wechsler-Bellevue and ACE tests with Number Anxiety, are presented in Table 4. The only correlations which are significant are those between grades and Number Anxiety.

Discussion

Several lines of evidence lead to the position that probably there is an overall dimension that may be termed Number Anxiety which differs from "general anxiety." The low correlations of the Taylor scale scores with the individual Number Anxiety items as well as with the total of the latter is the first piece of evidence. The second is the result of the cluster analysis which sharply separates Number Anxiety from the two "general anxiety" clusters. The third line of evidence is the highly significant changes which occurred in galvanic skin resistance for "number anxious" persons on the arithmetic instructions and arithmetic subtest of the Wechsler, changes which did not occur for the "general anxiety" group.

Although the identification of this overall dimension seems fairly secure, the actual unidimensionality of Number Anxiety is not assured by the study at hand. In fact, there is reason to believe that there are at least two subfactors in the surface factor isolated here, the two brought to light by the "little factor analysis" of Number Anxiety items. A further analysis of the three latter items in the matrix of 17 items is in progress. Also, an analysis of further items designed to ferret out Number Anxiety is being conducted with the 40 *S*s who constituted the original subsample.

Nothing in the present study yields bases for determining the etiology of Number Anxiety. From previous thought and spontaneous remarks by *S*s of the subsample, a number of hypotheses have been formulated as to the origin of this type of anxiety which *S*s find very handicapping. And very tentatively in the continuing research we are putting some of these hypotheses to the test. However, more work is needed to identify just exactly what Number Anxiety is and what differentiates it from other kinds of anxiety before much exact research can be done in determining etiology. One hypothesis which is a partially etiological one, that Number Anxiety is merely the result of low intelligence, seems to be set aside by the research already done, in that neither the Wechsler nor ACE showed any significant relation to Number

Anxiety. That actual accomplishment in mathematics is associated with lack of Number Anxiety is demonstrated by the present research, though how much causation is involved is a question, for one may cause the other, vice versa, or both may be the resultant of other factors.

One caution must be voiced in connection with the changes in GSR and Number Anxiety. These variables are associated, but the GSR cannot be strictly considered to be a measure of Number Anxiety, for other conditions of stress have also been shown to have GSR correlates. Apparently the Wechsler Verbal Scale is not sufficiently stressful to generate significant GSR changes except in connection with "number anxious" persons and Arithmetic.

Conclusions

All three of the proposed hypotheses were substantially supported by the research described here.

1. Number Anxiety does appear to be a separate (composite) factor from "general anxiety," although the correlation of .33 between the Taylor Scale and the three Number Anxiety items indicates that some relation probably holds between these two gross syndromes.

2. Number Anxiety does not seem to be related to general intelligence as measured by either the Wechsler-Bellevue or ACE scales. The supposition that Q scores on the ACE are correlated negatively with Number Anxiety must wait further research, for the r of $-.25$ does not reach significance.

3. Persons with high Number Anxiety do tend to make lower mathematics grades, as indicated by the r's of $-.44$ for the 704 *Ss* and $-.51$ for the 40 subsample group between Number Anxiety and grades in Mathematics 105.

References

1. Buss, A. A. A follow-up item analysis of the Taylor Manifest Anxiety Scale. *J. clin. Psychol.*, 1955, **11**, 409–410.
2. Cattell, R. B. The chief invariant psychological and psycho-physical functional unities found by P-technique. *J. clin. Psychol.*, 1955, **11**, 319–343.
3. Gordon, E., & Sarason, S. B. The relationship between "test anxiety" and other anxieties. *J. Pers.* 1955, **23**, 317–323.

UNIVERSITY OF MICHIGAN LIBRARIES

4. Holzinger, K. Jr., & Harman, H. H. *Factor Analysis, a synthesis of factorial methods*. Chicago: Univer. of Chicago Press, 1941.
5. Hoyt, D. P., & Magoon, T. M. A validation study of the Taylor Manifest Anxiety Scale. *J. clin. Psychol.*, 1954, **10**, 357–361.
6. Klein, E. Psychoanalytic aspects of school problems. In A. Freud, H. Hartmann, & E. Kris (Eds.), *Psychoanalytic study of the child*. Vol. III–IV. New York: International Universities Press, 1949. Pp. 369–390.
7. Kramer, G. A. *The effect of certain factors in the verbal arithmetic problem upon children's success in the solution*. Baltimore: Johns Hopkins Press, 1933.
8. Levy, D. M. *Maternal overprotection*. New York: Columbia Univer. Press, 1947.
9. Nichols, R. C., & Daroge, T. An electronic circuit for the measurement of the Galvanic Skin Response. *Amer. J. Psychol.*, 1955, **58**, 455–462.
10. O'Connor, J. P., Lorr, M., & Stafford, J. W. Some patterns of manifest anxiety. *J. clin. Psychol.*, 1956, **12**, 160–163.
11. Plank, Emma N., & Plank, R. Emotional components in arithmetical learning as seen through autobiographies. In A. Freud, H. Hartmann, & E. Kris (Eds.), *Psychoanalytic study of the child*. Vol. IX. New York: International Universities Press, 1952.
12. Schonell, F. J. *Diagnosis of individual difficulties in arithmetic*. Edinburgh: Oliver and Boyd, 1937.
13. Silverman, R. E., & Blitz, B. Learning and two kinds of anxiety. *J. abnorm. soc. Psychol.*, 1956, **52**, 301–303.
14. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.
15. Wechsler, D. *The measurement of adult intelligence*. Baltimore: Williams and Wilkins, 1944.

Received February 13, 1957.

THE RELATION OF DEPARTURE TIME AND RETENTION TO ACADEMIC ACHIEVEMENT

ABRAM M. BARCH

Michigan State University[1]

For over three decades there has been a slight but constant interest in the study of persistence and its relationship to achievement, especially academic achievement. Despite this interest, several promising suggestions and studies have not been followed up. Thornton (**5**) hypothesized that the more the test of persistence approximates the school situation, the higher would be the correlation between persistence and school success. Briggs and Johnson (**2**) found that a persistence measure obtained in the achievement situation itself (namely, time to complete an achievement test given without a time limit) was related to achievement when general verbal ability was taken into account. French (**3**) developed a short test of persistence which he found resulted in an improvement in prediction of academic success above that obtained by conventional aptitude tests. His test consisted of 30 number series problems given with the instructions that some problems were insoluble and were to be so marked and was scored in terms of amount of time spent on the test.

Problem

The present study had two objectives. The first objective was to repeat the Briggs-Johnson study in a situation where more information would be available about the aptitudes of the students and the characteristics of the achievement test items and to see if a better estimate of the predictive value of their persistence measure could not be obtained. The second objective was to determine the effect on achievement test score of requiring all students to remain in the testing situation for a period of time distinctly longer than most students would stay without restraint.

The assumption behind the second objective was that the merely spending more time in the testing situation, irrespective

[1] The author wishes to thank the Board of Examiners of Michigan State University for the assistance rendered in carrying out this study.

of the reason(s) for spending this additional time, might be the factor responsible for increased achievement. The particular method selected for retaining students in the testing situation beyond their usual departure time was chosen to interfere least with the normal procedure of achievement testing. In addition to this aspect of minimal interference, the reason given for the detention was chosen so as to be irrevelant (or neutral) to any given person's level of persistence or achievement aspiration level.

Procedure

Achievement test

At Michigan State University, all students are required to take a one-year sequence in natural science unless they have had courses in that area at other institutions or unless they can pass special qualifying examinations in the area. The third term of this course is concerned with the atomic-molecular theory and the related concept of energy. It includes laboratory exercises on gas laws, atomic theory and electricity, and the nature of matter.

The final examination of the course, new each term, is prepared by a group of course instructors working in conjunction with members of the Board of Examiners. The test consists of 120 multiple-choice five-alternative items. The test items are intended to measure comprehension rather than knowledge of specifics. Classification of test items in the examination used in this study (spring of 1955) according to categories based on the *Taxonomy of Educational Objectives* (**1**) indicated that the test contained 8 items on knowledge of specifics, 38 items on knowledge of ways and means of dealing with specifics, 45 items concerned with ability to make use of materials or ideas through interpretation, 14 items concerned with the use of abstractions in particular and concrete situations, and 15 items concerned with the analysis of relationships.[2]

Procedure

The final examination was given to all students in the course at the same time in the same building. Instructional classes were grouped for examination purposes. The number of students per room when tested varied but averaged about 60.

[2] Clarence Nelson performed the classifications.

Subjects were students in two sets of six classrooms each so selected that no student was in an examination room with a proctor who was also his natural science instructor. In all rooms distribution of the tests was supposed to begin at 10:10 A.M.

In six of the rooms there was no change from normal testing procedure. The proctors merely recorded the departure time of students in units of five minutes each from the time the tests were distributed. The first student left just less than an hour after the test began. The testing session was terminated at 12:30 or shortly thereafter. The students in these rooms formed the initial basic group.

In the other six testing rooms, the proctors were requested to write the following on the blackboard: Please remain seated until 12:00. This is necessary to reduce disturbance by those finishing early.

The proctors were also asked to call the students' attention to the notice before beginning the test administration. Three of the instructors failed to follow the full instructions and the students in these classrooms were omitted from further analysis. The remaining students formed the initial experimental group.

Analysis

In order to obtain a more homogeneous group with respect to educational background, only those students who had entered the University in fall, 1955, and who had taken the entire orientation test battery were retained for further study. This group included a small proportion of transfer students in addition to freshmen. The remaining basic groups were composed of 190 men and 108 women. The remaining experimental groups were composed of 134 men and 48 women. Men and women were treated separately because intercorrelations between tests and achievement measures had been found to be related to sex at M.S.U. with the women usually showing higher intercorrelations than men.

The aptitude tests in the orientation test battery that were utilized in this study included the L and Q portions of the ACE (1940 version), a locally prepared Reading Test (RT) and a locally prepared Arithmetic Proficiency Test (A). The intercorrelations of these tests and the predictive value of the RT and the A for a large group of M.S.U. freshmen has previously been published (**4**). The RT was the best single predictor of first term grade point average.

Results

Significance of departure time

Inspection of the correlational plots relating departure time to the other measures indicated that the relationships could be assumed to be linear. This is in contrast to the results obtained by Briggs-Johnson who generally found a U-shaped trend between departure time and achievement. In the present study, the longer the students stayed the better was their achievement score.

Table 1 gives the product moment intercorrelations for natural science final (NS), departure time (DT), RT, A, L and Q. Values for men appear above the diagonal; values for women are below the diagonal. The relationship between DT and NS is low. This was expected since verbal ability was expected to obscure the relationship between these two measures; i.e., leaving at a given time after the beginning of the test means one thing when done by a person with good verbal ability and something else when done by a person with low verbal ability.

It will be noted that DT correlated negatively with all other measures except NS and A for both men and women. The positive relationship with A and the tendency for A and DT to include similar variance in multiple combinations was somewhat puzzling. Further investigation suggested the apparent reason. A was given with such a liberal time limit as to be essentially an untimed test, and the scores on this test apparently included some variance due to this untimed aspect.

TABLE 1

INTERCORRELATIONS OF ORIENTATION TEST SCORES, DEPARTURE TIME AND NATURAL SCIENCE FINAL SCORE[a]

	1	2	3	4	5	6
1. Natural Science Final		.16	.24	.43	.35	.23
2. Departure Time	.14		−.25	.10	−.18	−.16
3. M.S.U. Reading Test	.44	−.22		.25	.63	.29
4. M.S.U. Arithmetic Test	.51	.13	.35		.27	.45
5. L scale, ACE	.50	−.10	.77	.38		.44
6. Q scale, ACE	.32	−.15	.47	.68	.45	

[a] Men (190) above diagonal; women (108) below diagonal.

TABLE 2

PARTIAL CORRELATION COEFFICIENTS BETWEEN NATURAL SCIENCE FINAL SCORE AND DEPARTURE TIME

Comparisons	Men (190)	Women (108)
$r^{a}_{12.3}$	.23**	.27**
$r_{12.4}$	.13	.09
$r_{12.34}$	.18*	.20*
$r_{12.5}$	.24**	.22*
$r_{12.6}$	.20**	.20*
$r_{12.56}$	.25**	.24*

* .05 level of significance.

** .01 level of significance.

[a] See Table I for names of variables.

Table 2 gives the first and second order partial correlation coefficients between the various predictor measures and NS. All partials are significantly different from zero at the .05 level or better except the first order partials with A as the partialled variable.

Another way of evaluating the importance of departure time as a measure for predicting academic achievement is to determine the extent to which it increases the magnitude of a multiple correlation. Table 3 lists the R's obtained for combinations of the predictor variables with and without DT when NS and grade point average for first M.S.U. year were the criterion variables.

Some loss in cases occurred in the group with grade point average as the criterion. Excluded from this group were students who did not return to school the next fall term and students who had less than 32 credits or more than 63 credits at the end of their first M.S.U. year. The remaining groups were composed of 146 men and 90 women.

In all comparisons except one, addition of DT to the predictor variables significantly increased the size of R at the .05 level or better. The one exception which involved the prediction of NS for women using RT and A as predictors was borderline to significance at the .05 level. The increase in predictive variance in all cases was small, but it must be remembered that departure time is essentially one response. The results suggest the value of obtaining additional measures of this type for use in prediction of academic achievement.

TABLE 3

MULTIPLE CORRELATIONS PREDICTING NATURAL SCIENCE FINAL SCORE OR FIRST YEAR GRADE POINT AVERAGE WITH DEPARTURE TIME EXCLUDED AND INCLUDED

Criterion Variable	Predictor Variables	*R*'s for Men	*R*'s for Women
Natural Science	3[a], 4	.45	.58
Final Score	2, 3, 4	.48	.60
(Men = 190)	5, 6	.36	.51
(Women = 108)	2, 5, 6	.43	.55
First Year Grade	3, 4	.36	.73
Point Average	2, 3, 4	.40	.77
(Men = 146)	5, 6	.39	.53
(Women = 90)	2, 5, 6	.45	.57

[a] See Table 1 for names of variables.

Sex differences

The mean NS score for men (basic group) was 61.4 and that for women (basic group) was 56.5. The difference was significant at the .01 level. Men and women were not significantly different in DT, RT, A or L. The men scored higher (.01 level) on the Q scale than did the women.

The point of interest in these results is that the sexes did not differ in terms of test persistence. It is possible that this finding is due to the nature of the test and that somewhat different results would obtain in a situation where the women would be superior in test achievement.

Effect of Retention

The results of the basic groups indicated that 74% of the women and 85% of the men could have been expected to leave by 12:00, the end of the retention period for the experimental groups. An analysis of covariance was performed for men and women separately using RT score as the control variable. The *F* in both comparisons failed to reach the .05 level. Thus, no evidence was obtained in support of the hypothesis that merely spending more time in the testing situation would lead to better achievement. It should be recalled that the manner of retention was designed to provide minimal interference with normal testing procedure and the reason given for the retention was neutral with respect to a

given person's level of persistence or aspiration level. The next question to be asked is whether instructions stressing the importance of taking more time on the test will affect achievement.

Conclusions

1. Persistence as measured by time taken by college students to complete an achievement test was found to be significantly related to test achievement and grade point average when verbal ability was partialled out.

2. The persistence measure significantly improved multiple regression predictions for academic achievement.

3. No difference was found between men and women in test persistence.

4. Retaining students in the testing situation with instructions which were neutral with respect to achievement had no affect on achievement.

References

1. Bloom, B. S. (Ed.) *Taxonomy of educational objectives*. New York: Longmans, Green, 1955.
2. Briggs, Arvella, & Johnson, D. M. A note on the relation between persistence and achievement on the final examination. *J. educ. Psychol.*, 1942, **33**, 623–627.
3. French, J. W. The validity of a persistence test. *Psychometrika*, 1948, **13**, 271–277.
4. Jackson, R. A. Prediction of academic success of college freshmen. *J educ. Psychol.*, 1955, **46**, 296–301.
5. Thornton, G. R. The use of tests of persistence in the prediction of scholastic achievement. *J. educ. Psychol.*, 1941, **32**, 266–273.

Received February 26, 1957.

DIFFERENTIATING CHARACTERISTICS OF A GROUP OF STUDENTS HAVING PSYCHOLOGICAL PROBLEMS

GEORGE DOUGLAS MAYO

Naval Air Technical Training Command, Memphis[1]

Virtually everyone associated with an educational program has had occasion to observe students with psychological problems and, frequently, to inquire whether or not the school situation is in some degree responsible for the condition observed in the student. Rotter (**5,** pp. 416–423) has discussed this matter in an astute manner and has attempted to treat it within the general framework of his social learning theory. In addition to the orientation described in the following paragraphs, a section of the present paper has been devoted to consideration of the findings of the study in terms of Hull's system (**3**) as adapted to emotional behavior by Brown and Farber (**1**).

It is important from the practical point of view to know whether psychological problems tend to be more closely associated with factors in the immediate environment or more closely associated with factors more remote from the current situation. If it is known that maladjustment is associated primarily with factors in the immediate environment, it may be possible to change these factors in such a way as to alleviate the condition. If not, it is necessary to look to other, probably more difficult, remedial procedures.

The study was conducted within this general framework of situational versus nonsituational causation of psychological problems and disorders. But no claim can be made that it provides definitive results concerning this general problem. In this connection, it should be noted that it was necessary to use available variables which could be classified as situational or nonsituational instead of following the more desirable procedure of sampling a universe of situational and nonsituational variables. The study does, however, identify and cross validate certain variables which

[1] The opinions expressed are those of the writer and are not necessarily shared by the Department of the Navy.

differentiate to a statistically significant extent navy trainees requiring assistance with psychological problems from those not requiring such assistance.

For purposes of the study, situational variables were defined as those primarily in the present environment, while nonsituational variables were defined as those essentially independent of the present environment.

Procedure

The study was conducted at a training center for naval aviation technicians having approximately 5500 navy and marine enlisted men under instruction throughout the period of the study. This period was from July 1953 through June 1954.

As shown in Tables 1, 2, and 3, the seven situational variables consisted of the schools in which the students were enrolled and measures of success achieved by the students in the schools. The nine nonsituational variables consisted of aptitude test scores, overachievement-underachievement in a previous navy school, educational level, age, service membership (i.e., Navy or Marine Corps), sex membership, petty officer status versus nonpetty officer status, and recruits versus men returned from the fleet for training.

The study is actually two almost indentical studies, one for the period of the last half of 1953, and the other for the first half of 1954. It was decided in advance that the null hypothesis would be rejected only in those instances in which differences were significant at less than the .05 level in the case of both samples. It might be argued that the same end could be achieved by selecting a more exacting confidence level for the rejection of the null hypothesis, but this is not quite true. We know that conditions change to some extent both within the schools and with respect to the students assigned to the schools over a period of time. In the present study, we were interested in whether the changes that occurred over a period of several months would affect the significant differences observed in the initial sample, as well as in the stability of the initial findings from the standpoint of sampling error. Selecting a more rigorous confidence level would satisfy only the second of these two interests.

The psychological problems groups consisted of all men who were so classified by the medical staff at the naval dispensary

to which they reported, either on their own initiative or were sent by the school which they were attending. Cases which were quite serious were transferred immediately to a naval hospital near the training center. Less serious cases were treated at the dispensary as outpatients. Thus, the psychological problems groups included students having a wide range of problems and behavioral disorders. The total number in the group for the last half of 1953 was 170, and for the first half of 1954, 162. This does not represent a decrease, since the population of the training center underwent a proportional decrease.

It was considered likely that certain schools were contributing a disproportionate share to the psychological problems groups. This proved to be true as shown in Table 1. This information, supplemented by the fact that number of students, aptitude test scores, and failure rates vary from school to school made it advisable to stratify the two samples of normal students with which the psychological problems groups were compared. The stratification resulted in the normal groups containing the same proportion of students from each school that the psychological problems group for the same time period contained. Having arrived at the proportion of the normal group that would be drawn from each school, the appropriate number was drawn by putting the IBM cards of all men attending the school during the period in service number order and drawing every nth card. If this stratification had not been accomplished it would have been difficult to state whether observed differences between the psychological problems groups and normal groups were due simply to school membership rather than to the variables per se. Two samples of 250 normal students were formed in this way, one from students attending the schools during the last half of 1953 and the other from students attending the schools during the first half of 1954. For purposes of the study, all students were considered to be normal if they had not been seen by the medical staff for psychological reasons. In the comparisons in which it was appropriate to compute means, the critical ratio statistic was used. In those instances in which the comparison of numbers of subjects was appropriate, chi square was computed.

Results and Discussion

The results of the study are shown in Tables 1, 2, and 3. Tables 2 and 3 contain all of the comparisons in which the chi-square

TABLE 1
SCHOOL MEMBERSHIP AND PSYCHOLOGICAL PROBLEMS

Aviation School	1953 Sample			1954 Sample		
	Observed	Expected[a]	Chi Square	Observed	Expected	Chi Square
Engines	13	32.8	14.8**	14	27.2	7.7**
Structures and Hydraulics	24	20.6	00.6	30	17.3	10.4**
Electronics	114	84.8	20.1**	100	81.0	8.9**
Training devices	11	10.2	00.1	10	9.2	0.1

[a] Expected frequencies are based upon the proportion of students in each school to total students at the training center.
** Significant at .01 level.

statistic and critical ratio statistic respectively were used, with the exception of school membership. In both of these tables the variables have been classified as being primarily situational or nonsituational.

School membership which is viewed as a situational variable is considered in Table 1. In the case of both the 1953 and 1954 samples, the aviation electronics school contributed a disproportionately large number of students to the psychological problems groups, and the aviation engines school contributed a disproportionately low number. It is interesting to note that the electronics course is more abstract and more difficult than the engines course. Furthermore, despite the assignment of students with higher aptitude to electronics, the course is failed by a considerably larger proportion of students than is the aviation engines course.

Table 2 indicates that both the 1953 and 1954 psychological problems groups contained a significantly larger proportion of academic failures than did the normal groups. From the data presented, however, one cannot be sure whether school failure tended to cause the higher incidence of psychological problems or whether a tendency toward psychological problems caused the observed higher failure rate. In addition to membership in certain schools and academic failure, a difference, significant at the .01 level, was found in the 1953 comparison on another variable classified as situational. This variable was school grades, as shown in Table 3. The same comparison for 1954 did not reach the standard set for statistical significance but closely approached it by

TABLE 2

COMPARISON OF PSYCHOLOGICAL PROBLEMS GROUP WITH NORMAL GROUP ON CERTAIN VARIABLES USING THE CHI SQUARE STATISTIC

Comparisons	1953 Groups					1954 Groups				
	Psychological Problems		Normal		Chi Square	Psychological Problems		Normal		Chi Square
	o	e	o	e		o	e	o	e	
Situational										
Academic failures	29	19	18	28	9.90**	24	17	20	27	4.79*
Other students	141	151	232	222		138	145	230	223	
Academic setbacks	84	79	112	117	.87	72	76	122	118	.75
Other students	86	91	138	133		90	86	128	132	
Nonsituational										
Petty Officers	13	11	15	17	.44	12	12	18	18	.01
Nonrated men	157	159	235	233		150	150	232	232	
Recruits	135	140	210	205	1.45	110	125	208	193	13.07**
Fleet returnees	35	30	40	45		52	37	42	57	
Navy men	110	123	193	180	7.86**	96	107	177	166	5.86*
Marines	60	47	57	70		66	55	73	84	
Male	164	166	245	243	.43	152	157	247	242	7.96**
Female	6	7	5	4		10	5	3	8	

* Significant at .05 level.
** Significant at .01 level.

reaching the .06 level of confidence. It should be noted that adequate data concerning school grades were not available for the students who failed to complete the schools. Hence the comparison on the basis of school grades includes only 67 of the 170 students in the 1953 psychological problems group and 42 of the 167 students in the 1954 psychological problems group.

The comparison involving the one remaining situational variable, academic setback, as shown in Table 2 did not yield a significant difference. The term "setback" refers to repeating a unit, usually two weeks in length, because of unsatisfactory progress in the unit.

Considering now the variables classified as nonsituational and on which significant differences were found in both the 1953 and 1954 samples, Table 2 indicates that the Marine Corps tended to contribute men to the psychological problems groups to a greater extent than did the Navy. Two additional variables classified as

TABLE 3

COMPARISON OF PSYCHOLOGICAL PROBLEMS GROUP WITH NORMAL GROUP ON CERTAIN VARIABLES USING THE CRITICAL RATIO STATISTIC

Comparisons	1953 Groups					1954 Groups				
	Psychological Problems		Normal		CR	Psychological Problems		Normal		CR
	Mean	*SD*	Mean	*SD*		Mean	*SD*	Mean	*SD*	
Situational										
School grades	73.15	5.00	75.25	5.64	2.89**	73.71	5.41	75.47	5.51	1.91
Nonsituational										
Marine aptitude battery, composite score	55.76	5.99	58.56	5.85	2.43*	55.14	5.88	56.53	6.96	1.25
Navy aptitude battery, composite score	57.16	5.18	58.65	4.71	2.46*	54.76	5.90	56.00	6.65	1.57
Over-underachievement index	98.97	5.23	100.14	4.91	2.00*	98.9	6.11	101.84	5.72	4.59**
Civilian education	11.21	1.21	11.82	1.27	4.77**	10.88	1.40	11.52	1.22	4.71**
Age	20.92	2.48	21.54	2.56	2.42*	20.94	3.67	21.35	2.90	1.18

* Significant at .05 level.
** Significant at .01 level.

nonsituational in Table 3 met the standard set for significance in both the 1953 and 1954 samples. The first of these was over-achievement–underachievement in a previous service school, the psychological problems groups tending toward underachievement. The second variable was years of civilian education, the mean of the psychological problems groups being less than that of the normal groups.

The remaining nonsituational variables did not meet the condition set for significance for the study. These included aptitude scores, age, sex membership, petty officer versus nonpetty officer status and recruit versus fleet returnee. It should be noted in this connection that these results do not necessarily indicate that no differences exist on these variables. The conditions of the study were not sufficiently rigorous with respect to numbers and perhaps with respect to other considerations to justify the statement that any existing differences were given an adequate chance to be observed. It is believed, however, that one can have reasonable confidence in the stability of the differences that satisfied the standard set for the study for the *rejection* of the null hypothesis.

Theoretical Considerations

Empirical data frequently are better understood in the light of theory. Similarly, improvement of theory depends to a major extent upon its exposure to empirical data. One of the more highly developed theories of behavior is that of Hull (**3**). Brown and Farber (**1**) have added certain constructs to Hull's system with a view to a more adequate handling of emotional behavior. Further, they have illustrated how their constructs fit into the larger system by considering phenomena commonly ascribed to frustration, in terms of their proposed method of conceptualization. The present study clearly does not lend itself to the rigorous procedures laid down by Hull for checking a segment of a theory and revising it in accordance with the results of the experiment. Nevertheless, it is of interest to see to what extent Hull's system, as expanded by Brown and Farber, gives unity to the observed functional relationships and facilitates understanding of them. Reference to Brown and Farber's diagram summarizing the relationships between antecedent conditions, intervening variables, and consequent conditions (p. 492) will be helpful in the following discussion. This diagram has been reproduced by Lambert also (**4**, p. 66.).

One of Brown and Farber's key explanatory concepts is frustration. They use the term to connote a hypothetical state or condition of the individual, as opposed to a frustrating event or external barrier. The term is viewed as an intervening variable between other constructs or intervening variables which ultimately are anchored both at the antecedent end and consequent or response end by observable conditions. In the present study, the antecedent conditions may be thought of as including the 16 variables on which the normal group and the psychological problems group were compared. As noted previously, six of these antecedent conditions were found to be significantly related to the consequent condition, that is, behavior which resulted in the individual being classified in the psychological problems group.

It is assumed that successful completion of the aviation schools was an important intermediate goal or means to an end for virtually all students. The technical schools are a usual route to advancement and status in the various aviation occupations. Moreover, most of the students had asked to be sent to the school in which they were enrolled in preference to some other aviation school.

Adapting the theory to the present data the intervening variable, "drive," D_O, is envisioned as a function of the desire to complete the school, and D_C as a function of some other desire or need. More than one of these competing drives, D_C, may have to be taken into account in a given situation. This concept of drive is a departure from Brown and Farber's diagram on two counts. First, drive is described in terms of desired goals instead of deprivation or noxious stimulation. Second, in the Brown and Farber diagram, D without subscripts is used and defined as "relevant drive." Hull has defined D as "strength of dominant primary drive operative in the primary motivation to action after the formation of the habit involved." The use of D_O and D_C insures that the appropriate drives will be applied to the habits involved, that is, to habit strengths having the same subscripts as the drives.

The intervening variables ${}_sH_{R_O}$ and ${}_sH_{R_C}$ represent habit strength of an ongoing response tendency directed toward achieving the goal of school completion, and habit strength of an ongoing response tendency directed toward achieving some other goal, respectively. When ${}_sH_{R_O}$ and ${}_sH_{R_C}$ are multiplied by their respective drives two additional constructs emerge, ${}_sE_{R_O}$ and ${}_sE_{R_C}$, which represent initial excitatory strength of the response tendency directed toward school completion and inital excitatory strength of the response tendency directed toward a goal other than school completion. The intervening variable "competitive inhibitory tendency," I_R, is envisioned as a function of blocking and nonreinforcement of effort to achieve the goal of successful school completion. "Frustration," F, the key intervening variable, results from the competition or incompatability of ${}_sE_{R_O}$ with both I_R and ${}_sE_{R_C}$.

At this point it may be well to see how well the six variables on which significant differences were found fit into the theoretical structure described thus far. Membership in the electronics school is associated with high I_R values and membership in the engines schools with lower I_R values since the electronics school has a much higher failure rate than does the engine school. Similarly, after drawing a sample of normal students stratified to contain the same proportions from each school as were contained in the psychological problems group, we confirm our notion that school failure is significantly related to membership in the psychological problems group. This would result in a larger value of I_R in the individuals that ultimately made up the psychological problems

group than in the normal group. These larger I_R values would more nearly equal the value of ${}_sE_{R_O}$ in the psychological problems group than in the normal group, a condition which, according to the theory, increases the value of F. The results concerning the antecedent conditions, educational level, and overachievement-underachievement in a previous navy school can be explained in a similar way. It is reasonable that the behaviors which underlie both of these variables are related to success or failure in the aviation schools. The explanation of the final variable on which a significant difference was found, membership in the Marine Corps versus membership in the Navy, is seen in the larger value of F in the case of marines due to competition between the incompatible response tendencies ${}_sE_{R_O}$ and ${}_sE_{R_C}$. The present study provides no data on the subject, but there is considerable informal evidence to indicate that marines join the corps with the goal of becoming a member of a conventional marine combat unit. Many are not volunteers for aviation, as is true in the case of navy men. This results in a higher D_C value for marines, which, when multiplied by its appropriate habit strength, ${}_sH_{R_C}$, results in a larger value of ${}_sE_{R_C}$ for marines than for navy men. Similarly, the conditions that result in a larger value of ${}_sE_{R_C}$ may also result in a somewhat smaller value of ${}_sE_{R_O}$ than that of navy men. For some individuals ${}_sE_{R_C}$ may exceed ${}_sE_{R_O}$. The behavior resulting from this condition under different circumstances will be considered later. In any event, the competing values of ${}_sE_{R_O}$ and ${}_sE_{R_C}$ would be expected to be more nearly equal in the case of the marine group than the navy group, and hence a larger value of F would be expected.

Brown and Farber hold that F has two major effects. First, it increases the general level of motivation. This effect is integrated into Hull's theory by considering F as an "irrelevant drive," for which Hull has presented formulas. In the present adaptation of the theory "effective drive strength," $\overline{D}_O$, results from the combination of the drive to complete the school, D_O, and the "irrelevant drive due to frustration," $\dot{D}$. When ${}_sH_{R_O}$ is multiplied by this new drive strength, $\overline{D}_O$, the construct, "final excitatory strength of the response tendency directed toward completion of the school," ${}_sE_{R_O}'$, emerges. $\dot{D}$ also combines with D_C, which, through the same process as that just described, results in ${}_sE_{R_C}'$, which is the "final excitatory strength of the response tendency directed toward a goal other than school completion."

The second major effect of F is to produce unique internal stimuli, S_F. These stimuli may have innately associated response tendencies, ${}_{s_F}U_R$, or they may acquire new associative connections, ${}_{s_F}H_R$, through processes of learning. These response tendencies are largely maladaptive and may bear little rational relationship to the objective toward which the thwarted drive was directed. These response tendencies, when multiplied by S_F, result in the final construct in the system, ${}_{s_F}E_R$, "strength of excitatory tendency associated with frustration."

At this point we have three final response tendencies, ${}_sE_{R_O}'$, ${}_sE_{R_C}'$, and ${}_{s_F}E_R$. Assuming threshold strength, the response which will occur depends upon the relative strength of the three final response tendencies. If ${}_sE_{R_O}'$ has the greatest magnitude, behavior directed toward completion of the school will occur. This may take the form of additional study, seeking assistance from instructors or classmates outside of class, and/or other adaptive behavior related to successful school completion. If ${}_sE_{R_C}'$ is the strongest of the three final response tendencies, the other two will be inhibited, and behavior directed toward a goal other than school completion will occur. In the case of marines who joined the corps with a desire to be a member of a conventional marine combat unit, it may take the form of behaving in a manner calculated to obtain a transfer to such a unit. If ${}_{s_F}E_R$ is greater than the other final response tendencies, the response will tend to be maladaptive. In a military organization, perhaps to a greater degree than in a civilian setting, this type of behavior is likely to result in a referral for psychological problems. Unlike Dollard et al. (2), Brown and Farber do not specify precisely which response, e.g., aggression, withdrawal, or primitivization, will occur. The theory does include the expectations that the presence of frustration will often result in the intensification of response, that reactions followed by a reduction in frustration will be reinforced, and that responses associated with frustration in one situation will tend to appear in other thwarting situations.

The above discussion obviously is no more than exploratory in adapting the Hull, Brown and Farber theory to the type of data described. Despite the numerous limitations, including explanation after the fact and an absence of quantification, it is interesting to note the apparent ease with which the observed relationships can be handled by the theory. It raises the important question of

whether the method employed in the present study could be adapted to the task of providing some of the empirical data required for detailed systematic development of the Brown and Farber theory of frustration. Perhaps this question could be answered by one or two studies in which the antecedent conditions were selected in accordance with the best information currently available, and measured as carefully as possible. The consequent conditions used in the study are already suited to the theory, since in its initial form it does not attempt to specify the precise reaction to frustration that will occur.

Summary

Two groups of navy and marine trainees having psychological problems were compared with two groups of normal trainees on the basis of 16 variables. Significant differences between both psychological problems groups and their corresponding normal groups were found on six of the comparisons. Students of aviation electronics tended to have psychological problems in greater numbers than did other students. The reverse was true in the case of men studying aviation engines. Academic failure was associated with membership in the psychological problems groups. The three comparisons just mentioned were thought of as being primarily situational in nature.

The remaining three comparisons on which significant differences were observed were classified as nonsituational variables. These were: overachievement-underachievement in a previous navy school, years of civilian education, and membership in the Navy versus the Marine Corps. Thus no consistent tendency for maladjustment to be more closely associated with situational variables or with nonsituational variables was found.

The findings were discussed in terms of Hull's theory, as adapted to emotional behavior by Brown and Farber.

References

1. Brown, J. S., & Farber, I. E. Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, **48**, 465–493.
2. Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. R. *Frustration and aggression*. New Haven: Yale Univer. Press, 1939.
3. Hull, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.

4. Lambert, W. W. Stimulus-response contiguity and reinforcement theory in social psychology. In G. Lindzey, (Ed.), *Handbook of social psychology*. Cambridge: Addison-Wesley, 1954.
5. Rotter, J. B. *Social learning and clinical psychology*. New York: Prentice-Hall, 1954.

Received March 28, 1957.

COMPARATIVE EFFECTIVENESS OF TELEVISED AND LARGE CLASSES AND OF SMALL SECTIONS[1]

LAURENCE SIEGEL AND F. G. MACOMBER

Miami University, Ohio

College enrollments for the next decade are expected to rise sharply with an anticipated shortage of well-qualified college teachers in most subject areas. These conditions will place additional pressures upon public institutions to increase class size. The development of educational television, and particularly of relatively low-cost closed-circuit systems, raises the interesting possibility of teaching large groups of students while overcoming some of the visual difficulties inherent in large class instruction.

Miami University is engaged in a research program directed toward the comparison of the relative effectiveness of large group instructional procedures with smaller conventional-type classes of approximately 30 to 35 students. Two of the general approaches to large group instruction under investigation are designated LC (large classes with direct visual contact between student and instructor) and TV (closed-circuit television) classes. This paper summarizes the major results obtained during the Spring Semester, 1956. Three problem areas were investigated:

1. Achievement as a function of assignment to experimental or control sections;
2. Students' attitudes about the instructor and course as a function of section assignment;
3. Students' end-of-the-semester attitudes about having received TV or LC instruction rather than conventional instruction.

Procedure

Subjects

The first two problem areas under investigation required the comparison of results obtained from experimental and control sections of the eight undergraduate courses listed in Table 1.

[1] This research was conducted under a grant to Miami University from the Fund for the Advancement of Education.

Each experimental section was taught either as a large class (primarily lecture) or by means of closed-circuit television. The control sections were taught by conventional procedures. Both experimental and control sections were taught by the same instructor and no known biasing factor was known to be operative as a determinant of student registration for the experimental or control section. Nevertheless, the relatively small sizes of the control groups prohibited sole dependence upon random assignment of students to balance the effects of extraneous variables. Consequently, the experimental and control groups for each course were deliberately equated on the basis of scores on the Cooperative English Test (Form T), Cooperative Mathematics Achievement Test (Form P), and total score on the American Council on Education Examination (1948 edition), as well as grade-point-average for the previous semester.

The matching procedure required that some of the students actually enrolled in the experimental and control sections of each course be excluded from the specific experimental and control groups under investigation. All decisions about the constituency of the experimental and control groups were made prior to the collection of research data. Neither the instructors nor the students were informed regarding the actual composition of the groups providing data for the Experimental Study in Instructional Procedures. The total enrollments in each course and the numbers of subjects included in the study are summarized in Table 1.

A full description of the matching data is given elsewhere (2). It is sufficient here to indicate that none of the mean differences between the experimental and control groups in any course on any of the matching variables was statistically significant. Equating of the experimental and control groups on the critical variables of ACE and grade-point-average was further supported by the lack of statistical significance of the chi squares between pairs of score distributions on these two variables. Furthermore, both groups in each of the eight courses were demonstrated to be equated both with respect to proportional sex distributions and proportional distributions by class standing.

Teaching Procedures

The specific instructional procedures within each course were not standardized. Rather, instructors were encouraged to develop

TABLE 1

TOTAL ENROLLMENT AND NUMBERS OF EXPERIMENTAL SUBJECTS IN EACH COURSE

Course Title	Number of Sections			Number of Students Enrolled in Sections[a]			Number of Subjects from Each Section[b]		
	TV	LC	Control	TV	LC	Control	TV	LC	Control
Foundations of Human Behavior	1		2[c]	103		74	97		62
Principles of Human Physiology	1		1	96		31	90		30
Introductory Sociology	1		4[d]	141		144	131		129
Human Biology	1	1		183	178		183	177	
Business and Government		1	1		119	35		105	30
Composition and Literature		2	2[e]		100	50		37	37
Essentials of Modn. Geography		3	3[e]		213	93		88	85
Elementary Psychology		1	1		78	30		77	30

[a] Enrollments in multiple sections were divided approximately equally between each of the sections.

[b] These are the numbers of students from each section actually providing data for the Study.

[c] Multiple sections taught by the same instructor.

[d] Multiple sections taught by different instructors, each of whom rotated TV presentations.

[e] Multiple sections taught by different instructors, each of whom also taught an experimental section.

those procedures which seemed best suited to the instructional situation, and which were best calculated to achieve the objectives of the course. Thus there were considerable differences in the manner in which the four televised courses were conducted. The Foundations of Human Behavior class met for two weekly periods of 90 minutes each; Sociology and Human Biology were convened for three weekly periods of 50 minutes each. The students in the experimental Physiology section met for three 50-minute televised presentations and for a two-hour laboratory period each week. The nature of the televised presentations also differed in the four

courses. In both Biology and Physiology, the general method was that of lecture supplemented by numerous demonstrations and training aids. Five different members of the department participated in the televised presentations of Sociology. These presentations were approximately 30 minutes in length with the balance of the 50-minute period devoted to a discussion in each receiving room under the leadership of a regular staff member. In Foundations of Human Behavior, the presentations took a variety of forms, but with approximately one-third of each class period given over to discussion under the leadership of graduate assistants who served as monitors in each of the receiving rooms.

Similarly, there was no standard form of procedure in the large classes not utilizing television as a medium of presentation. The LC sections of English followed essentially the same procedure as that designed for the control sections except that only half as many written themes were required, and, of course, individual student participation was considerably limited by the size of the experimental groups. In Geography, the students received two lectures a week with a third period devoted to small group discussion. The remaining two LC courses were given as three 50-minute lectures per week.

The control sections of all courses were conducted as discussion groups with lectures and demonstrations inserted by the instructor whenever he thought it desirable.

This diversity of approach was believed desirable even though it prohibited direct comparisons between courses. Forcing the instructors to adhere to a specific and inflexible type of presentation in both their experimental and control sections would have prevented them from capitalizing upon the advantages inherent in the specific classroom situation and would have emphasized some of the deficiencies.

Evaluative Instruments

Two classes of instruments provided data for this series of investigations: achievement tests and attitudinal measures. These two classes of instruments are described separately below.

"Achievement" was operationally defined for the present sequence of studies as performance on the objective portions of course examinations administered for the purpose of assigning final grades. It is recognized that such examinations typically

measure only a single dimension of achievement, i.e., subject-matter knowledge. Thus the present data are not germane to such important course objectives as "synthesis," "problem solution" and "critical thinking." The possibility of leakage of information about the test was overcome by simultaneous administration of each examination to the experimental and control sections of each course with one exception. (Simultaneous administration in Biology was impossible because the experimental and control sections were taught during successive semesters.) Estimates of reliability of each test (Kuder-Richardson Formula 21) ranged from .55 to .93 with a median of .88. To the extent that the assumptions implicit in this computation were not met, the resultant values are probably underestimates.

Four Thurstone-type scales were included in the attitude battery administered near the end of the semester to students in experimental and control sections. These scales yield quantifiable information about (*a*) students' appraisal of the effectiveness of the instructor in presenting material and handling the class (the I scale); (*b*) students' appraisal of the course content (the C scale); (*c*) students' attitudes about having received televised rather than conventional instruction (the TV scale); and (*d*) students' attitudes about having received large class instruction rather than conventional instruction (the LC scale). Each of these scales provided a nine-point continuum with 5.0 as the neutral position. Student judges were requested to rate the items included in a preliminary pool, and items were selected for inclusion in the final forms of these instruments on the basis of low Q value and the median of the distribution of judgments.

Split-half reliabilities (corrected for length of test) based on responses of a sample of 300 respondents were .91 for the I scale and .92 for the C scale. The intercorrelation between these two attitude scales was .52. The corrected split-half reliability for the TV scale was .89 and for the LC scale was .92 in a sample of 100 students.

Results

Conclusions based on the results obtained from a single semester's instruction are subject to the influence of a variety of sampling errors. Only four TV and four LC courses were investigated. Conclusions based on these data might not apply to other courses or even to the same content taught by other instructors. Also,

TABLE 2

COMPARATIVE ACHIEVEMENT IN EXPERIMENTAL AND CONTROL SECTIONS

Course	TV Phase				Course	LC Phase			
	Section	Achievement		*t* ratio		Section	Achievement		*t* ratio
		M	*SD*				M	*SD*	
Foundations of Human Behavior	TV	123.76	16.03	0.25	Business and Government	LC	70.03	8.58	0.70
	Control	123.11	15.05			Control	71.28	7.65	
Introductory Sociology	TV	110.09	20.91	0.24	Composition and Literature	LC	28.94	5.59	1.33
	Control	109.53	16.90			Control	30.78	4.74	
Physiology	TV	208.92	31.33	0.66	Essentials of Geography	LC	79.83	10.59	0.83
	Control	204.59	26.44			Control	78.54	9.56	
Human Biology	TV	253.97	15.21	3.00*	Elementary Psychology	LC	109.45	15.93	1.77
	Control	248.66	15.76			Control	102.52	21.77	

* Significant at better than the .001 level.

results obtained from the present samples of students may not generalize to other samples enrolled in the same courses chiefly because of the relatively small size of some of the control groups.

Achievement

The distributions of achievement test scores in the experimental and control sections as well as appropriate *t* ratios are summarized in Table 2. The null hypothesis is not refuted for any course except Biology, i.e., with the exception of this one course, comparable achievement resulted from experimental and control presentations.

Interpretation of the statistically significant finding in the case of Biology is confounded by the fact that the experimental and control sections of this course were given during successive semesters rather than simultaneously. Although a definitive interpretation of the results obtained for Biology is impossible at the present time, it is apparent that neither TV nor LC presentations adversely affected achievement as defined by subject matter examinations. This finding for TV courses is in agreement with results reported earlier by the Pennsylvania State University (**1**).

The absence of differential achievement in experimental and control sections does not negate the possibility of differential achievement as a function of academic ability in interaction with

TABLE 3
COMPARATIVE ACHIEVEMENT AS A DUAL FUNCTION OF SECTION ASSIGNMENT AND LEVEL OF ABILITY

Course	Section	High Ability Subgroup				Low Ability Subgroup			
		Achievement Test			*t*-ratio	Achievement Test			*t* ratio
		M	*SD*	*N*[a]		M	*SD*	*N*[a]	
Foundations of Human Behavior	TV	126.28	29.74	43	0.34	115.81	16.06	43	1.37
	Control	128.38	13.92	26		120.96	14.77	29	
Introductory Sociology	TV	117.80	13.13	55	0.97	103.04	14.67	46	0.12
	Control	115.52	12.60	67		102.70	9.54	40	
Physiology	TV	221.70	29.34	37	0.41	197.87	29.20	39	0.25
	Control	217.73	19.27	11		195.76	27.28	17	
Human Biology	TV	258.03	13.81	95	2.20[c]	249.56	18.84	86	2.61[c]
	Control	253.14	16.50	98		241.45	17.12	68	
Business and Government	LC	70.86	8.06	49	0.81	69.02	6.93	51	0.31
	Control	72.86	8.14	14		69.80	8.89	15	
Composition and Literature	LC	32.43	2.79	14	0.06	26.73	5.57	22	2.09[b]
	Control	32.33	5.29	15		29.73	3.95	22	
Essentials of Modn. Geography	LC	83.81	9.19	42	1.14	76.11	10.49	43	0.22
	Control	81.59	8.32	41		75.63	9.80	44	
Elementary Psychology	LC	113.02	13.98	44	1.03	104.53	17.15	32	1.75
	Control	108.61	17.69	18		92.55	24.06	11	

[a] Fluctuations from total *N* in the sample caused by unavailability of ACE scores or absence from final examination.
[b] $p < .05$. [c] $p < .02$.

section assignment. In order to investigate the possible existence of such interaction, the students within each experimental and control group were dichotomized with respect to ability on the basis of percentile conversion of total ACE store. The cutting point for this dichotomization was the fiftieth percentile. Comparative achievement of the "high ability" subgroups in the experimental and control sections and of the "low ability" subgroups in these sections is summarized in Table 3.

The data presented in Table 3 support the conclusion that level of ability does not interact with assignment to a TV section or its control counterpart as a dual determinant of achievement. A similar conclusion applies to all LC courses with the exception of Composition and Literature. The statistically significant difference for the "low ability" students in this one instance is the result of both a mean difference and a variance difference. It might have resulted from the unreliability of the final examination or from sampling

errors. If replicated with a more reliable criterion and larger samples, it would indicate that this is one course wherein the less able students profit more from small class instruction than from LC presentations. The high ability students in this course, however, performed equally well under both classroom conditions.

Students' Attitudes About the Instructor and Course

It is desirable to camouflage the identity of specific courses and instructors for presentation of attitudinal data. These data are presented for courses designated by a letter rather than a title. The order in which the courses are hereafter listed differs from the order of listing in the previous discussions.

Group data summarizing the results obtained from administration of the I scale and the C scale in the experimental and control sections are summarized in Table 5. In courses taught by multiple instructors, the mean scores on both scales were obtained by pooling the data across instructors.

It is apparent from the data exhibited in Table 4 that TV instructors were rated at least as favorably by students in their television sections as by students in their control sections. Two of the four LC instructors, however, received significantly less favorable ratings in the experimental section than in the control section.

Referring again to Table 4, there is a pronounced tendency for students in the control sections to rate the course content more favorably than students in either the TV or LC sections, although the obtained difference is statistically significant in only three of the seven courses for which data were available. Apparently, instruction given simultaneously to large groups of students (either by TV or in large classes) is accomplished at some sacrifice of the students' favorability to the course even though it does not affect their examination performance. Although the data are not amenable to direct comparison of TV and LC instruction, they lead to the suspicion that this deficiency may be more marked in the case of LC instruction than it is for TV instruction.

Attitudes About TV and LC Instruction

Group responses to the TV and LC scales are summarized in Table 5. The fiducial limits extending the mean allow for an absolute interpretation of this value in terms of favorability or

TABLE 4

SUMMARY OF I-SCALE AND C-SCALE RATINGS IN EXPERIMENTAL AND CONTROL SECTIONS OF EACH COURSE

Course	Section	I Scale[a]			C Scale[b]		
		M	*SD*	*t*-ratio	M	*SD*	*t* ratio
A	TV	2.80	1.00	—	3.08	0.50	—
	Control[b]	—	—		—	—	
B	TV	4.56	1.74	1.35	4.50	1.24	2.27*
	Control	4.10	1.40		4.00	0.91	
C	TV	4.16	0.88	2.86**	4.39	1.26	0.94
	Control	4.56	1.24		4.23	1.26	
D	TV	2.94	1.02	0.11	3.43	0.88	0.94
	Control	3.14	1.00		3.28	0.85	
E	LC	4.64	1.16	2.14*	3.91	1.07	3.18**
	Control	4.00	1.48		3.21	0.86	
F	LC	3.58	1.06	2.67**	3.83	1.07	2.88**
	Control	3.10	1.16		3.37	0.89	
G	LC	3.42	1.08	0.77	3.97	1.10	1.26
	Control	3.22	0.98		3.68	0.83	
H	LC	3.60	0.80	0.00	3.93	0.95	1.90
	Control	3.62	0.40		3.50	0.95	

[a] Scores below 5.0 denote ratings as "more effective than average;" the lower the mean score, the more favorable the rating.

[b] Data not available.

* Significant between the .05 and .01 levels.

** Significant at better than the .01 level.

TABLE 5

SUMMARY OF TV AND LC SCALE RESPONSES

TV Scale				LC Scale			
Course	M	*SD*	Fiducial Limits*	Course	M	*SD*	Fiducial Limits*
A	4.91	0.83	±0.14	E	5.74	0.72	±0.16
B	6.19	0.81	±0.18	F	5.45	0.84	±0.19
C	5.42	0.99	±0.20	G	5.35	0.81	±0.26
D	4.76	0.94	±0.20	H	6.03	0.64	±0.13

* Fiducial limits of the mean calculated for the .05 level of confidence.

unfavorability to TV and LC instruction (taking 5.0 as a neutral position).

Two conclusions follow from these data:

1. It is impossible to generalize on students' attitudes about TV

instruction. Students in one course preferred TV instruction to conventional instruction (Course D); those in another course considered the two types of presentation to be of approximately equal effectiveness (Course A). However, students in two of the TV courses (B and C) felt that television was inferior to conventional instruction.

2. Students in all LC courses would have preferred assignment to a small section. The degree of this unfavorability, however, varied from slight (Courses F and G) to considerable (Course H).

The apparent variability in TV and LC attitudes from course to course is indicative of the fact that factors other than section assignment play a role in conditioning students' attitudes about the mode of instruction. The potential influence of two such factors (students' ratings of the instructor and academic ability) was investigated. Correlations between both I scale responses and ACE total score with responses to the TV and LC scales are exhibited in Table 6. Note that significant correlations involving the ACE are indicative of an inverse relationship, i.e., the higher the ACE score, the less favorable the TV or LC attitude.

It is apparent that if any ability subgroup is unfavorable to TV or LC instruction, it is more likely to be the high ability students than the low ability students. However, the inverse relationship between academic ability and attitudes toward the mode of presentation may be overcome by factors which are not yet entirely clear. One such mitigating factor is probably student perception about the instructor. There is a marked tendency for academic

TABLE 6

CORRELATIONS BETWEEN I-SCALE RESPONSES AND ACE TOTAL SCORE WITH ATTITUDE TOWARD EXPERIMENTAL TYPES OF INSTRUCTION

TV Phase			LC Phase		
Course	I Scale	ACE	Course	I Scale	ACE
A	.25[b]	.00	E	.32[b]	.06
B	.10	.32[b]	F	.13	.02
C	.17	.23[a]	G	.10	.38[a]
D	.36[b]	.14	H	.24[a]	.04

[a] r exceeds 1.96 times standard error of r of 0.00 ($p < .05$).

[b] r exceeds 2.58 times standard error of r of 0.00 ($p < .01$).

ability to correlate inversely with TV and LC attitudes in those courses where ratings of the instructor do *not* correlate with these attitudes.

Conclusions

1. Acquisition of subject-matter knowledge as measured by objectivity retest performance was not adversely affected by assignment to large (experimental) section (TV or LC) rather than to a small (control) section.

2. Academic ability did not interact with type of section assignment (experimental or control) as a complex determinant of achievement in any of the televised courses. A similar conclusion applies also to the LC courses under investigation with the possible exception of Composition and Literature.

3. Student ratings of the instructor (I scale) were not adversely affected by television. The only statistically significant difference between ratings by the TV and control sections occurred in one course wherein the instructor was rated more favorably by his TV section. Ratings of the instructor did tend, however, to be less favorable in LC sections than in control sections. The mean difference was statistically significant in two of the experimental courses.

4. Student ratings of the course (C scale) were less favorable in both TV and LC sections than in the control sections. This deficiency appears to be even more serious in LC sections than in TV sections (although data amenable to a direct comparison between these two types of presentation were not available).

5. Students' attitudes about TV instruction were not uniformly held across courses. In contrast, students in all LC courses were unfavorable to the mode of instruction when compared to small class instruction.

6. When academic ability was related at all to students' TV and LC attitudes, the direction of this relationship was inverse, i.e., high ability students were less favorable than low ability students. However, the fact that such a relationship was not present in all courses suggests the likelihood that other factors interact with academic ability in this regard. One such interacting factor is probably the students' attitudes about the effectiveness of the instructor.

References

1. Carpenter, C. R., & Greenhill, L. P. *An investigation of television for teaching university courses.* University Park: Pennsylvania State Univer., 1956.
2. Macomber, F. G., & Siegel, L. *Progress report of the experimental study in instructional procedures.* Oxford: Miami Univer. 1956.

Received April 25, 1957.
Early Publication.

BOOK REVIEW

D. B. KLEIN. *Mental Hygiene: A Survey of Personality Disorders and Mental Health.* New York: Henry Holt, 1956. Pp. 654.

This is the first revised edition of a book which appeared in 1954. It follows the same general plan and seeks to follow the committee which considered ways of improving the teaching of psychology, especially in emphasizing intellectual development rather than an applied interest. This book, therefore, is not merely a book on personal adjustment but deals more fundamentally with principles of mental hygiene.

Part One, Chapters 1 and 2, discusses the nature and scope of mental hygiene and gives a summary of its history and problems. Part Two, Chapters 3 and 4, deals with the nature and kinds of mental disorder. Part Three, Chapters 5 through 9, takes up the most important problem of the prevention of mental disorder. Part Four, Chapters 10 through 18, on the promotion of mental health, includes discussions of the home and the personality; developmental problems from infancy through adulthood; the dynamics of conscience; fixations; coping with reality; problems of repression, conflicts, frustration, anxiety, emancipation, and self-control; economic problems, security, morale, and the aged; two popular forms of idolatry—grades and competitive success—and the contribution of the right kind of religion to mental health.

"There are many ways to achieve personality integration, many ways to keep an even keel through the storms of life." There are many individual differences. There are other and perhaps better ways than the "hazardous trail of the go-getter." Recognition of the proper values in life may help us to avoid the "exaggerated devotion to dollar-chasing as the chief virtue of daily life." It may help us to a healthier life than that of "relentless competition and power-seeking and self-aggrandizement."

Although the book is not written primarily from the point of view of applied psychology, but rather from that of fundamental principles, the reader can find many practical applications in many parts of the book. Especially, he should be helped to a clearer understanding of psychology, particularly in its relation to mental health; and he can do so without being lost in technical language and unnecessary, complicated jargon, unverified assumptions, and

reified elements, that are prevalent today and often pass under the name of psychology. It is generally clear and simple in presentation, and a glossary (pp. 627–642) adds distinctly to the usefulness of the book.

A few minor adverse criticisms seem relatively unimportant in comparison with the general excellence of the book. It is sufficiently comprehensive for a semester college course and can be understood by the beginning student in psychology. It will stand easily with the few really excellent books in its field. A good index (pp. 643–654) concludes the book.

A. S. EDWARDS

The University of Georgia

An extraordinarily comprehensive, scholarly, and systematic treatment of personality and motivation measurement.....

PERSONALITY AND MOTIVATION STRUCTURE AND MEASUREMENT

by Raymond B. Cattell

Research Professor of Psychology, University of Illinois
Director, Laboratory of Personality Assessment and Group Behavior

This brilliant study by Professor Cattell makes available to applied and experimental psychologists the new resources of factor analytic measurement. It presents a particular theoretical system, complete to the point justified by research findings gathered intensively over the last twelve years (since the publication in 1946 of Cattell's *Description and Measurement of Personality*) and, more discursively, over thirty years. In addition, it delineates for the busy university teacher and student the results of research which is either new or has heretofore been reported in fragmentary form in widely scattered writings. An invaluable work for the serious student of personality.

WORLD BOOK COMPANY, Yonkers-on-Hudson, New York

Again Available

SYMPOSIUM: PSYCHOANALYSIS AS SEEN BY ANALYZED PSYCHOLOGISTS

This well-known symposium, which first appeared in the *Journal of Abnormal and Social Psychology* in 1940, has now been reprinted. The contributors are Edwin G. Boring, Carney Landis, J. F. Brown, Raymond R. Willoughby, Percival M. Symonds, Henry A. Murray, Else Frenkel-Brunswik, David Shakow, Donald V. McGranahan, and Austin B. Wood. Comments on the symposium are made by Hanns Sachs and Franz Alexander. 160 pages.

Price, $2.00

Order from

American Psychological Association

1333 Sixteenth Street N. W. **Washington 6, D. C.**